Location Choice, Linkages and the Spatial Economy: Essays on Theory, Evidence and Heterodox Assessment

David Stephan Bieri

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Planning, Governance & Globalization

Paul L. Knox, Chair

Casey J. Dawkins

Heike Mayer

Bradford F. Mills

12 August 2010

Blacksburg, Virginia

Keywords:

Regional linkages, nonmarket interactions, spatial structure, qualitative and quantitative discourse

Location Choice, Linkages and the Spatial Economy

Essays on Theory, Evidence and Heterodox Assessment

David Stephan Bieri

Abstract

The essays in this dissertation represent theoretical and empirical contributions to urban economics and regional science, focusing on the growing importance of nonmarket interactions. There is increasing evidence that the process of globalization is rendering the world "spiky" rather than "flat". Nonmarket interactions, such as knowledge spillovers, innovation or amenity-based externalities, play a central role in this process. As economic activity is not evenly spread across space, a detailed understanding of the economic linkages between regions is key to the design of effective public policy. This is particularly important in the context of economic linkages between regions or cities, highlighting the key adjustment mechanisms – via both market and nonmarket transactions – and their long-run implications for incomes, the cost of living, and the spatial distribution of population. Both the neoclassically-grounded field of urban economics and the rapidly expanding New Economic Geography (NEG) literature pioneered by Krugman offer a variety of models and (not infrequently competing) predictions about the factors and processes that shape the spatial structure of the economy. At the same time, the dialogue between qualitative and quantitative discourses in regional science has been marred by an increasingly embittered dispute over methodology. While acutely pronounced in economics, this development has re-shaped large parts of its sister disciplines as well, particularly sociology and geography. Across the board, proponents of quantitative science methodology increasingly likened themselves to their natural science counterparts, whereas qualitative methods had become the last bastion of "true social scientists". Today, these so-called "science wars" have rendered "qualitative" and "quantitative" analysis into almost mutually exclusive concepts.

Acknowledgements

The work in my dissertation reflects a long and winding intellectual journey that has its origins almost two decades ago, taking me quite literally from A to B – from my native Aarau in Switzerland via London, Durham and Basel to Blacksburg in the US. I began my academic career as an undergraduate in the geography department at the LSE in the early 1990s, before moving across to the economics department during the time when the "cultural turn" in economic geography was in full swing. This proved to be a formative experience indeed, which – from the privileged vantage point of both sides of the disciplinary divide – has since committed to me to the idea of transdisciplinarity.

While at the LSE, it was a chance encounter with Kurt Klappholz that sparked my fascination with the potential inconsistencies of the classical dichotomy and the neoclassical synthesis in macroeconomics, progressively moving my interests up the scale of spatial aggregation. In the years that followed, macro-level monetary phenomena and international finance would become the focal point of my work, first at Durham University and then in my professional life at Bankers' Trust in London and later in various capacities at the Bank for International Settlement in Basel.

Throughout this journey, my father proved to be my most trusted and challenging intellectual sparring partner whose insistence on the importance of regional phenomena, in particular the real effects of local spillovers and interregional transfers, I found increasingly difficult to counter. The more time I spent in financial markets, the more dissatisfied I became with simply rejecting the classical dichotomy on account of sticky prices. In addition, my casual observations during regular central banking missions to cities around the globe – from London to New York, Tokyo and Shanghai or from Riyadh to La Paz, Maputo, and Dakar – often afforded me deeper economic insights than did the macroeconomic briefs for the respective countries. Clearly, the mechanics of the urban economy contained key elements that permitted a more nuanced understanding of the economic realities of globalisation. While the "proximity to the patient" could not be a substitute for analysing macroeconomic aggregates, I had become convinced that the scope of my work needed to be re-focused by increasing its spatial resolution.

The decision to return to university for a PhD arose out of this desire to reconcile global (financial) trends with a deeper understanding of local developments. In this context, I am indebted to John Dooley for convincing me that the School of Public and International Affairs (SPIA) at Virginia Tech was the optimal locus for such a venture, and to Paul Knox for agreeing to guide me through this process as my dissertation supervisor. I am particularly grateful for Paul's indulging of my meandering at the early stages of my PhD, while diligently keeping me on the straight and narrow of a manageable dissertation time line. I hope that the breadth of the work in my dissertation is testimony to his encouragement to "be voracious", without jeopardising rigour and focus. I would also like to thank Paul for helping me to effortlessly straddle the transatlantic culture gap. I never felt like a "non-resident alien" in Blacksburg.

I would like to express my gratitude to my committee members, each of whom in his or her unique way was instrumental in opening up new intellectual horizons and in contributing to the synthesis that became my dissertation. Heike Mayer introduced me to all things local economic development and to the power of combining qualitative and quantitative research methods. I am also grateful for the numerous research opportunities on which I was able to collaborate with her. Casey Dawkins rekindled my interest in public finance and offered me glimpses of how I could leverage my past experience on new academic interests. Brad Mills invited me to re-discover the dismal science, but this time at a different level of spatial aggregation which gave me a new appreciation of how to reconcile general equilibrium models with localised mismatches and regional imbalances.

Special thanks are also due to Nick Kuminoff and Jaren Pope for sharing ideas, involving me in captivating research projects and for their patience, mentoring and unwavering support in numerous empirical trials and errors and for getting results on paper. Chris Parmeter has given me the confidence that lousy matrix algebra could be fixed, or at least that it could be mitigated by representative geometry, and that respectable applied econometrics need not be tied to commercial software packages. Patricia Nickel has re-acquainted me with critical theory and renewed my attraction to the philosophy of science. Peter Schaeffer has done all of the above and has given me a very special friendship and a sense of home away from home, not only because of his inimitable moité-moité fondues and Süessmostcrème.

The last four years would not have been the same without the motley crew of my fellow graduate students in SPIA, whose vast array of interests would be very hard to match in any PhD programme and whose camaraderie will be missed.

Finally, I am deeply thankful to my family for their continuous support throughout this entire journey. Without their love, generosity, vision and faith, I simply would not have been able to manage the more trying aspects of the last four years. My wife Anja has been an inspirational, patient and loyal friend, ever encouraging in the process, and I am looking forward to the next stage in my career with her at my side.

Attributions

Nick Kuminoff and Jaren Pope aided in the writing and research behind chapter 3 of this dissertation. However, the major substance of this chapter is the original work of the author of this dissertation who is also responsible for all the writing of this chapter.

Prof. Nicolai V. Kuminoff (currently in the Department of Economics at Arizona State University) was an assistant professor in the Department of Agricultural and Applied Economics at Virginia Tech during the author's graduate studies and contributed to chapter 3 in terms of discussing the scope, structure and content.

Prof. Jaren C. Pope (currently in the Department of Economics at Brigham Young University) was an assistant professor in the Department of Agricultural and Applied Economics at Virginia Tech during the author's graduate studies and contributed to chapter 3 in terms of discussing the scope, structure and content.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and overview

There is increasing evidence that the process of globalisation is rendering the world "spiky" rather than "flat".[1] Nonmarket interactions, such as knowledge spillovers, innovation or amenity-based externalities, play a central role in this process. As economic activity is not evenly spread across space, a detailed understanding of the economic linkages between regions is key to the design of effective public policy. This is particularly important in the context of economic linkages between regions or cities, highlighting the key adjustment mechanisms – via both market and nonmarket transactions – and their long-run implications for incomes, the cost of living, and the spatial distribution of population. Both the neoclassically-grounded field of urban economics and the rapidly expanding New Economic Geography (NEG) literature pioneered by Krugman (1991a,b, 1995) offer a variety of models and (not infrequently competing) predictions about the factors and processes that shape the spatial structure of the economy. In addition to a recent surge in empirical testing of NEG hypotheses, the most promising area of research in this field is the integration of these two hitherto separate strands of scientific inquiry in regional science into a unifying framework.

At the same time, the dialogue between qualitative and quantitative discourses in regional science has been marred by an increasingly embittered dispute over methodology. Despite the mutual quest for a unifying science, this row has not only deepened many of the cross-disciplinary divides, but it has also created gaping rifts within disciplines. For many fields in social science, the immediate post-World

[1] See Christopherson, Garretsen, and Martin (2008) for a recent discussion of the literature.

War II period marked indeed a time of “metamorphosis” and departure from disciplinary orthodoxy – a trend that was particularly backed by emerging quantitative methods. While acutely pronounced in economics, this development has re-shaped large parts of its sister disciplines as well, particularly sociology and geography. Across the board, proponents of quantitative science methodology increasingly likened themselves to their natural science counterparts, whereas qualitative methods had become the last bastion of “true social scientists”. Today, these so-called “science wars” have rendered “qualitative” and “quantitative” analysis into almost mutually exclusive concepts.

In this selection of essays, I aim contribute to a narrowing of the quantitative-qualitative divide in regional science by contributing to the literature in two main ways. Examining the impact of agglomeration effects and externality spill-overs, essays 2 and 3 in part I of my dissertation identify different channels through which the location decisions of households and firms are separately determined. These essays represent theoretical and empirical contributions in urban economics and regional science.

Second, in contrast to the quantitative nature of part I, the essays in part II engage in a critical exploration of the discourse in contemporary economic theory. In essay 4, I offer a heterodox assessment of contemporary mainstream economic theorising and argue that a reconsideration of knowledge and power within the larger project of socio-economics might be pivotal to developing a new vision in economic theory. Essay 5 is an excursion into the history of economic thought and revisits one of the oldest controversies in economic theory, the so-called “transformation problem” – deriving prices from values and providing a theory of profits as arising from surplus value. While mainstream economists have by and large come to dismiss the transformation problem as a trivial technical exercise, the issue has recently received renewed attention in heterodox economic theory. Particularly the relationship between values and relative price formation in financial markets has emerged in the wake of the recent financial crisis.

Part III of this collection of essays is more exploratory and investigative in nature as it outlines several key components of a future research agenda in some detail.

1.1 Regional linkages and nonmarket interactions

The aim of this section is to broadly outline recent developments in the urban economic and NEG literature with regard to understanding regional inequalities and the clustering of economic activity. Summarising key differences and similarities, table 1.1 provides a point of departure for a synthesis of these two strands of literature and proves the larger theoretical setting for my dissertation work.

Table 1.1: Key concepts in geographical economics

Concept	*Urban economics*	*New Economic Geography*
Scale	City, region	Region (market-size dependent)
Spatial structure	Monocentric city	Core-periphery
Spatial linkages	No	Yes
Nonmarket interactions	Yes	No
Agglomeration effects		
First nature:	geography, resource endowments	–
Second nature:	urbanisation externalities (Jacobs)	–
	localisation externalities (Marshall)	–
	–	pecuniary externalities
Technology*	CRS	IRS
Goods sector	Traded manufacturing goods, non-traded housing services and amenities	
Transportation cost	drives rent gradients	drives agglomeration
Labour market†	mobile	mobile (C), immobile (P)
Price level	exogenous	endogenous
Money sector	No	No
Government sector	Yes	No

Notes: "Geographical economics" is a broad field that commonly includes the two main strands of urban economics and New Economic Geography, occasionally also extending further outward to the dynamics of urban system change, which is the terrain of regional economics. * CRS: constant returns to scale, IRS: increasing returns to scale; † (C): core, (P): periphery.

The most obvious distinction between urban economics and economic geography lies in the respective difference of the relevant spatial unit of observation. While the former primarily focuses on intra- and inter-urban phenomena, the object of inquiry for the latter are the spatio-temporal aspects of regions. Methodologically, economic geography (and regional science for that matter) are grounded in traditional macroeconomics, largely relying on aggregates that are above the level of spatial resolution of the micro-based units, such as the firm or individual households. The work of urban economists, on the other hand, is principally rooted in these very micro foundations.

Other differences are perhaps more nuanced and subtle, particularly with regard to nonmarket transactions. In contrast to the well-known nonmarket phenomena in the urban literature, including Alfred Marshall's human-capital spillovers and Jane Jacobs's creativity-based innovation, the NEG literature illustrates how theory can explain cities and agglomeration without addressing any nonmarket interactions. Indeed, Krugman's (1991b) seminal work derives part of its appeal from the ability to explain economic agglomerations without resorting to nonmarket effects.[2]

As table 1.1 illustrates, a further difference between the models developed in economic geography and those of urban economics is the different role they assign to costs whose origin lie in the spatial economy. For instance, the workhorse model in NEG does not accommodate various congestion costs generated by the emergence of urban agglomerations and, similarly overlooks other agglomeration economies such as a better matching on labor markets, the proximity of intermediate inputs, and the existence of local knowledge spillovers. Despite some very recent efforts for a unified framework (e.g. Thisse, 2010), these two strands of regional science still remain largely unconnected.

Figure 1.1 translates the concepts in table 1.1 into a simplified graphical representation of the spatial economy, highlighting the most important regional linkages between individual economic actors, goods and services. This stylised framework provides the theoretical backdrop for most of the work in my dissertation which analyses specific key adjustment mechanisms between regions in more detail.[3]

The question of how nonmarket interactions shape the spatial economy – be it via knowledge spillovers, innovation and power relations or amenity-based externalities – constitutes the overarching theme across all the essays of my dissertation. The first essay in chapter 2 is entitled "Booming Bohemia?" and focusses on the lower part of figure 1.1, exploring the mechanics of regional growth and development by looking at the role of occupation-based human capital formation in the US high-tech industry. Furthermore, this essay also investigates the sensitivity of these effects to different levels of spatial aggregation.

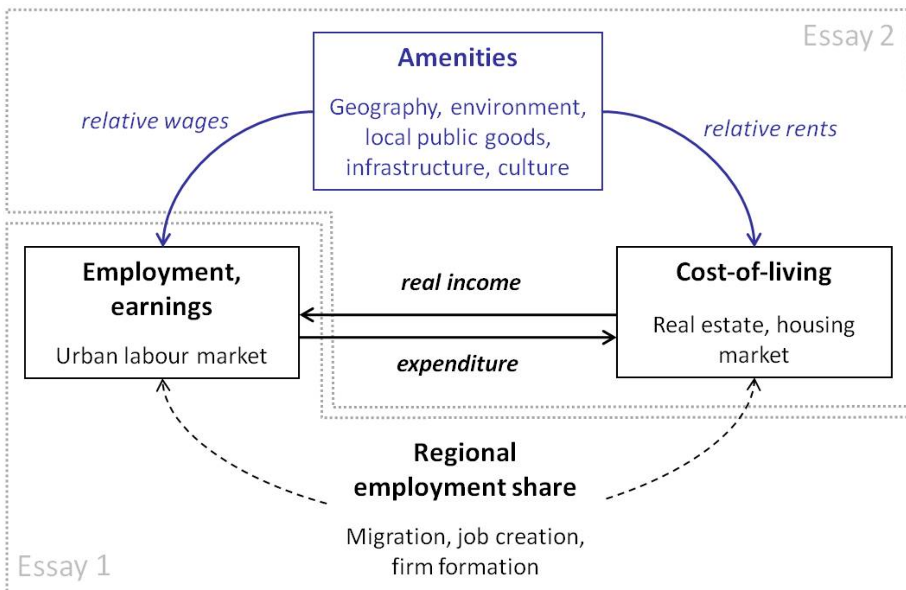
The second essay in chapter 3 examines the upper portion of figure 1.1 and is based on the observation that location-specific differences in wages and (land) rents should compensate for the differences in non-market characteristics, such as natural or cultural amenities that increase the attractiveness of a given locality. Entitled "Accounting for amenities", this work assesses the increasing regional importance of non-market goods, such as environmental amenities or local public goods, and proposes new estimates for personal consumption expenditures on quality of life in the US. These estimates complement the corresponding NIPA measures for private goods.

Part II of my dissertation is dedicated to work that is situated at the fringe of the mainstream, offering critical perspectives on elements that are either taken for granted or are explicitly assumed away in part I

[2] See Glaeser (2000) for a comprehensive overview of the increasing importance of nonmarket interactions in urban research.

[3] Figure 1.1 builds on the work of Overman, Rice, and Venables (2010) who explicitly develop a more formal model of regional linkages to analyse the impact of shocks, such as productivity improvements or housing supply increases, in one region on other regions

Figure 1.1: Regional markets and locational equilibrium



Notes: Essay 1 in chapter 2 is entitled "Booming Bohemia?" and explores the mechanics of regional growth and development by looking at the role of occupation-based human capital formation in the US high-tech industry. Essay 2 in chapter 3 is entitled "Accounting for amenities" and proposes new estimates for personal consumption expenditures on quality of life in the US. These estimates complement the corresponding NIPA measures for private goods.

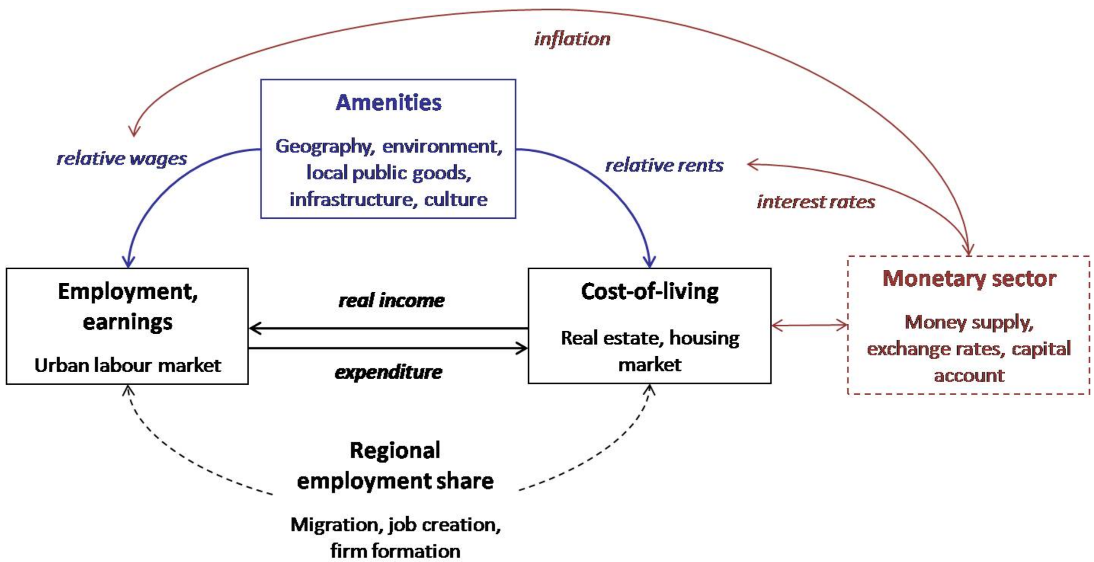
of my work. Specifically, this part of my work looks at . The third essay in chapter 4 provides an alternative perspective on the importance of (unobservable) linkages of a slightly different nature, namely those of power and knowledge. In this setting, knowledge is conceived to represent much more than just being a catch-all term for human-capital specific effects. In a sense, this work can be viewed as a complementary effort of describing the political economy behind the regional linkages described in part I. In a similar vein, perhaps, the fourth essay in chapter 5 revisits the old nexus of value theory and relative prices. While regional disparities in relative prices and their associated adjustment mechanisms constitute the main drivers of the spatial linkages within the stylised framework of figure 1.1, the transition from values to prices is implicity taken as given. Revisiting the puzzle of the transformation problem, the second essay in part II considers arguments – both old and new – why moving from values to prices might be more than just a technical exercise.

1.2 Extensions

Part III of my dissertation outlines specific elements of future work that evolved out of the main research presented in parts one and two. Chapter 6 is a direct offshoot of the research in chapter 2 and explores the spatial characteristics of urban specialisation and concentration in the United States. Similarly, chapters 7 and 8 are immediate extensions of the work in chapter 3, where the former presents some refinements of the theoretical framework and the latter seeks to embed the quality-of-life metrics within the larger context of the regional economy. Chapter 9 presents some current, stand-alone work in the area of applied econometrics and has some methodological relevance for the previous chapters. In particular, it ties into aspects for future work that are explicitly touched upon in chapter 2, but are also relevant for much of the other empirical work presented here. Be it households' or firms' location choices or agglomerations effects, the problem of endogeneity is of great concern to the applied researcher and the advances in instrumental variable methods illustrated in this chapter offer some remediation of the standard shortcomings of these techniques.

I would like to highlight two additional aspects of my research agenda here. The first element is essentially an extension of the work outlined in chapter 6 and relates to the notion of regional competitiveness. While regional competitiveness is usually analysed within a firm location setting, amenity-driven quality-of-life considerations for households – rather than productivity differentials – are becoming increasingly important in explaining spatial patterns of economic activity. My research aims to integrate

Figure 1.2: Regional goods markets and the monetary transmission mechanism



both elements of spatial sorting and to derive implications for regional policy.

The second aspect concerns the regional equivalent of the classical dichotomy between real and nominal variables. Monetary phenomena have traditionally been deemed beyond the scope of regional economic analysis. The standard view that globalisation implies the "end of geography for finance" further entrenches the neglect of regional price dynamics which have important consequences for local development in industrialised and developing economies alike. Be it the composition of the regional cost-of-living or the determination of real wages, money still is – always and everywhere – a local phenomenon with spatial components. Figure 1.2 presents an augmented version of the regional linkages discussed above by illustrating what type of monetary transmission mechanism might interact with real variables, which in turn would affect the original linkages.

Some of my other work on the international importance of these monetary linkages (Bieri, 2009) represents modest steps in this direction and my future research endeavours to explore these issues in more detail within the larger purview of both theoretical and applied regional science.

1.A References

BIERI, D. S. (2009): "Financial Stability, the Basel Process and the New Geography of Regulation," *Cambridge Journal of Regions, Economy and Society*, 2(2), 303–331.

CHRISTOPHERSON, S., J. H. GARRETSEN, AND R. MARTIN (2008): "The World Is Not Flat: Putting Globalisation in its Place," *Cambridge Journal of Regions, Economics and Society*, 1(3), 343–349.

GLAESER, E. L. (2000): "The Future of Urban Research: Nonmarket Interactions," *Brookings-Wharton Papers on Urban Affairs*, 2000(1), 101–138.

KRUGMAN, P. R. (1991a): *Geography and Trade*. The MIT Press, Cambridge, MA.

——— (1991b): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–499.

——— (1995): *Development, Geography and Economic Theory*, no. 6 in The Ohlin Lectures. The MIT Press, Cambridge, MA.

OVERMAN, H. G., P. RICE, AND A. J. VENABLES (2010): "Economic Linkages across Space," *Regional Studies*, 44(1), 17–33.

THISSE, J.-F. (2010): "Toward a Unified Theory of Economic Geography and Urban Economics," *Journal of Regional Science*, 50(1), 281–296.

Part I

Location Choice and Spatial Linkages

Chapter 2

Booming Bohemia? Evidence from the U.S. high-technology industry

Published in: *Industry and Innovation*, 2010, Vol. **17**(1): pp. 23–49, Special edition "Experience the Creativity", guest-edited by Richard Florida.

2.1 Abstract

This paper assesses the effect of Richard Florida's creative class concept on economic growth and development at two levels of spatial aggregation. First, I examine the dynamics of economic growth across U.S. metropolitan regions and investigate how they relate to regional specialisation and the concentration of talent in the high-tech industry. In addition to evidence of significant high-tech clusters, I identify important complementarities with regard to the interaction between the three Ts of regional development (talent, technology and tolerance) and regional growth dynamics. Using firm-level data, the regional analysis is then complemented by exploring the location of new high-technology plant openings and their relationship with university research and development (R&D) and the creative class. Specifically, I test the hypothesis that both university R&D *and* the presence of "creativity" generate spillovers which are captured locally in the form of new high-tech establishments, after controlling for important location factors such as local cost, demand and agglomeration economies. While the marginal impacts of increased R&D funding on county probability for new firm formation is modest, the mix of creativity and diversity – as proxied by the Florida measure – appears to be a key driver in the locational choice of new high-tech firms. Separate estimates indicate that these findings hold up across the major high-tech industries in the U.S.

Keywords: Regional economic development, firm location, high-tech industry, R&D, creative class, dis-

crete choice model
JEL Codes: R12, R39, C25

2.2 Introduction

Economists, geographers, economic developers, and policymakers have long been concerned with a comprehensive understanding of the mechanisms responsible for regional economic growth and development. Clear identification of the determinants of growth would allow for the design of more germane policies with greater potential to positively impact the economies of regions. The complex causes of urban growth and its associated spatial patterns of economic development are at the heart of regional science inquiry. In this context, researchers are particularly challenged by the question: do people move to jobs or do jobs move to people? One of the most promising approaches to tackle this issue is to relate urban growth directly to the economic geography of production (Storper and Scott, 2009).

Indeed, firms' location decisions often drive economic models of agglomeration. Broadly speaking, there are three classes of models of firm-level locational decision making: Firms value proximity (1) to customers in market access models, most notably in the New Economic Geography literature; (2) to other firms as sources of positive externalities in production externality models; and (3) to specialized inputs in natural advantage models. Special cases of production externalities are localization and urbanization externalities. These models have dominated the literature on firm location choice.

2.2.1 Firm location choice and localisation externalities

A vast theoretical and empirical literature shows that human capital accumulation and localised information externalities propel urbanization (cf. Lucas (1988)). In this context, it is frequently argued that advanced university R&D may also spur economic growth by fostering invention and spreading innovation across space. Indeed, there are many examples where high-technology (R&D-intensive) industry clusters seem to form around leading universities.[1]

In the United States, university-based urban growth centers are well known: the Silicon Valley in California, the Route 128 corridor around Boston, the Austin, TX metropolitan area, and the Raleigh-

[1] Most of the innovation literature uses knowledge generation and high-technology industries as synonymous. However, Hirsch-Kreinsen (2008) shows that "low-tech" industries are far from being characterised by firms that are not innovative. Innovations in sectors with little R&D-activity and those in R&D-intensive sectors are to a great extent interdependent, implying that innovation in low-tech industries should not be seen as a contradiction in terms.

Durham-Chapel Hill (Research Triangle) region of North Carolina. A series of empirical papers suggest that universities do matter as locational engines of regional economic growth as they provide anchors for local spatial externalities between university research and high technology innovative activity (e.g. Bania, Eberts, and Fogarty (1993); Anselin, Varga, and Acs (1997, 2000)).

More recently, discrete choice models have been deployed to assess the impact of knowledge spillovers on high-tech firm location choice. In particular, Woodward, Figueiredo, and Guimarães (2006, WFG hereafter) show that R&D expenditures at universities exert a positive, statistically significant influence on the decision to locate plants at the county level. Using a similar methodology, Andersson and Hellerstedt (2009) find evidence that supply-side conditions and market size support start-up activity in knowledge-intensive business services.

2.2.2 Complementary explanations? Universities and the "Creative Class"

At the same time, an alternative explanation for high-tech economic development and growth has been put forward by Richard Florida's notion of the "creative class" (Florida, 2002a,b). According to this concept, the underlying hypothesis is that the presence of creative "bohemians" in an area – rather than the ivory towers of academia in isolation – creates an "environment or milieu that attracts other types of high-talented or high-capital human beings" (Florida, 2002a, p.55).

Traditionally, the university's role in the economy is one of a relatively narrowly defined institutional locus for the transmission mechanism for industry research, inventions, patents and technology spin-offs. As such, universities are the "engines of innovation" whose success – and contribution to the economy as a whole – are hinged on the ability to commercialise research. Indeed, Braunerhjelm (2008) provides evidence from Sweden of a close correspondence between university research specialisation and regional industrial specialisation and productivity. In contrast to this view of institutional specialisation, Florida, Gates, Knudsen, and Stolarick (2006) suggest that the university's changing role within the larger regime of economic production is inextricably linked to the structural transformation of the knowledge-based economy. Specifically, they argue that " [...] the university plays a role not just in technology, but in all the Ts of economic development: technology, talent and tolerance." (Florida, Gates, Knudsen, and Stolarick, 2006, p.1).

The creative class approach has not only captured audiences at economic development agencies and other practitioners in the U.S. and abroad, but also sought academic legitimacy by providing empirical evidence of this driver of firm location choice (see section 2.3.1 for an overview of the recent empirical literature).

The contribution of this chapter to the literature is twofold. First, within a framework of geographic concentration, I investigate the extent to which the dynamics of urban growth in the U.S. are indeed influenced by the three Ts (talent, technology and tolerance) of economic development. Second, shifting my focus of analysis to industrial specialisation, I explore the location choices of new high-tech plant openings and their relationship with university R&D and a creative milieu. Given advances in empirical methods and better data, it is now possible to test location decisions as a choice among many alternatives. It is important to recognize that any establishment birth in a county involves a spatial selection based on relevant factors prevailing at the time. Specifically, I re-examine the hypothesis set out in Florida (2002a) with the WFG dataset and test if – in addition to university R&D – the presence of the "creative class" positively affects firms' decisions to locate new establishments in high-technology industries, after controlling for other important county characteristics.

The remainder of this chapter is structured as follows. Section 2.3 looks at the regional specialisation and concentration for the high-tech industry in the U.S. and investigates how firm location and the presence of knowledge-based spillovers from the three Ts of economic development influence regional economic performance. Section 2.4 then presents the discrete choice framework for firm location choice and section 2.5 discusses the data for the microdata-based analysis in some detail. Section 2.6 reports the empirical results and is followed by a brief conclusion in section 2.7.

2.3 Regional specialisation and concentration

A large amount of literature provides empirical evidence in support of Marshall (localisation) or Jacobs (urbanisation) theories regarding the specialization or diversity effects on the economic performance of regions. Beaudry and Schiffauerova (2009) provide a recent comprehensive survey of this literature, specifically focusing on the empirical evidence of the localisation versus urbanisation debate. To a large extent, measurement problems due to a high level of aggregation – both at the spatial and at the industry level – appear to be main causes for the lack of resolution in the debate. In part at least, my analysis is less prone to similar limitations since I am enhancing data with more commonly used higher-level aggregation (in this section) with high-resolution microdata (in section 2.5).

While it is clear that the mechanics of the transmission mechanism between knowledge and urban growth are complex, there is ample empirical evidence that industries where knowledge is a major input

Figure 2.1: Spatial distribution of high-tech concentration across US metropolitan areas, 2005

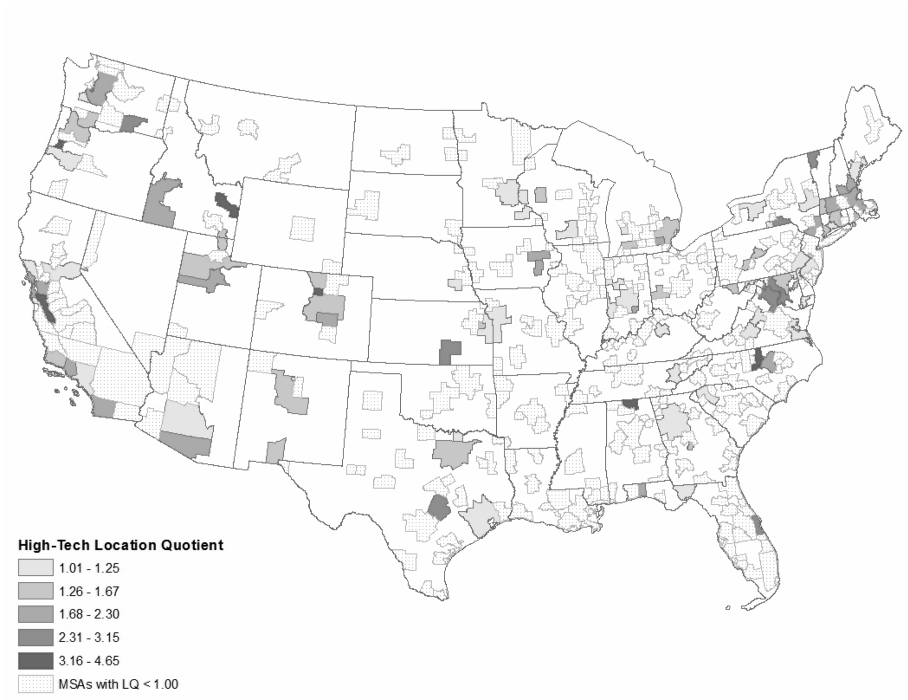
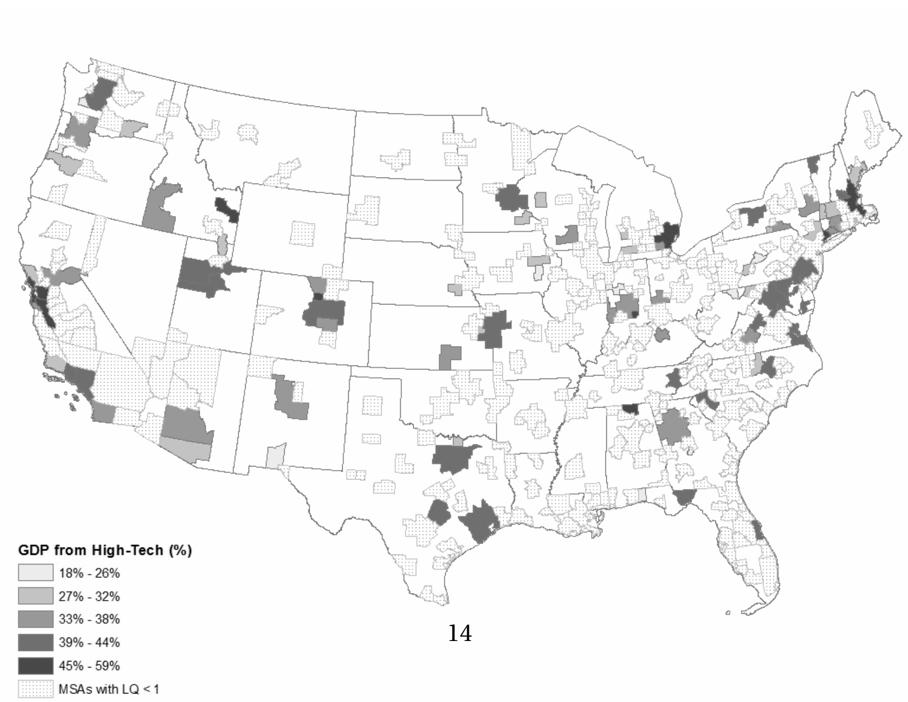


Figure 2.2: Metropolitan GDP from high-tech specialisation, 2005



Source: Author's calculations using a NAICS-based definition of high-tech industries following Hecker (2005) and Bureau of Economic Analysis (BEA) for the metropolitan GDP data (Panek, Baumgardner, and McCormick, 2007).

factor tend to cluster geographically. This localisation of knowledge-based firms is beyond what could be expected to happen by chance or what is primarily due to the natural geographical advantage of a given area (Ellison and Glaeser, 1997, 1999).

Acs, Audretsch, and Feldman (1992) substitute a measure of innovative activity for patents in the knowledge production function used in Jaffe, Trajtenberg, and Henderson (1993) and they find local university spillovers within 29 U.S. states. Similarly, using patents as a proxy for innovations, Audretsch and Feldman (1996) describe this geography of innovation in the U.S., linking the spatial clustering of industries to the presence of knowledge externalities. Complementary research by Acs and Audretsch (1988) provides evidence that innovative output is influenced by R&D and market characteristics, such as concentration and unionisation. They also find that these determinants have disparate effects for small and large firms. Ellison, Glaeser, and Kerr (2010) provide more recent evidence on the importance of agglomeration for intellectual spillovers among U.S. manufacturing firms. Using data from knowledge-intensive Canadian firms, Shearmur and Doloreux (2009) find evidence that different spatial patterns in the geography of innovation vary across industries.

In the context of the high-tech industry in the United States, the spatial characteristics of specialisation and concentration are illustrated in figures 2.1 and 2.2. A casual inspection of these two maps confirms that metropolitan high-tech concentration (as measured by the location quotient) and its specialisation counterpart (as proxied by the percentage of urban GDP from high-tech) are closely related.[2] While a variety of mechanisms could account for these spatial agglomeration patterns, different theories make different predictions about which industries should be coagglomerated in which regions. In the context of knowledge spillovers, we can broadly distinguish between three types of agglomeration economies in the context of innovation: (i) localised learning (Glaeser, 1999), (ii) product and process development cycles (Saxenian, 1994), and (iii) quality-based competition (Porter, 2000).

In this paper, I do not attempt to distinguish between different types of agglomeration effects, but assume that spillovers can be due to any one of these forces. However, localised learning is the main channel through which agglomeration economies are expected to operate in the present context, an intuitive argument that has been popularised by Florida (2002b).

[2] Chapter 6 explores spatial aspects of regional differences in industrial specialisation and concentration in more detail.

2.3.1 Growth and the three Ts of regional development

The major conjecture of the creative class hypothesis links urban development with novel combinations of knowledge and ideas. It suggests that certain occupations specialize in this task, that people in these occupations are drawn to areas providing a high quality of life, and that the essential development strategy is to create an environment that attracts and retains these workers. Specifically, Lee, Florida, and Acs (2004) find evidence that the creative class is strongly associated with new firm formation and high-tech specialization in metropolitan areas. Among the rapidly growing empirical literature that links the spatial distribution of the creative class to geographic disparities in economic growth, there is mounting evidence from Europe supporting the hypothesis of a positive impact of bohemians on regional economic development and entrepreneurship (see e.g. Boschma and Fritsch, 2009 and Falck, Fritsch, and Heblich, 2009).

The creative class theory thus posits that creative people are attracted to places most conducive to creative activity. While developed with urban areas in mind, this thesis may also be particularly relevant in rural areas, which lose much of their young talent as high school graduates leave for college, the armed forces, or "city lights". In this context, McGranahan and Wojan (2007) analyse recent developments in rural US counties, which focuses on natural amenities as quality of life indicators, and find support for the creative class thesis. Additional evidence is presented in Wojan, Lambert, and McGranahan (2007) who model the share of creative employment at the county level to explain differences in various measures of economic dynamism.[3]

However, there is a long-standing debate that questions the unequivocal link between the creative class and economic growth versus other models of (urban or regional) growth. See e.g. Hoyman and Faricy (2009) for a recent set of empirical tests and Blumenthal, Wolman, and Hill (2009) for alternative explanations in this context. At the heart of this debate lies the notion that different measures of human capital (such as education or the occupation-based creative class) appear to suggest different transmission mechanisms of how human capital drives growth. In an attempt to overcome this dispute, Florida, Mellander, and Stolarick (2008) propose a path model of regional development that formalises a general notion of regional growth that is somewhat heuristically based on the reinforcing interaction of the three Ts of economic development. The remainder of my analysis presented in this section is conducted in the

[3] McGranahan, Wojan, and Lambert (2010) apply the creative class theory to explain how some rural places may be economically dynamic even in a national context where growth depends on the novel combination of knowledge and ideas. Their rural variant of the creative class construct re-emphasizes outdoor amenities as an attractor of talent.

spirit of this model.[4]

Table 2.1: Top 10 MSAs by talent, technology and tolerance, 2000–2006

	Talent		*Technology*		*Tolerance*	GDP p.c.	Growth
	Creative class	Univ.§	LQ	GDP from HT	Bohemians	($, 2006)	(2000–06)
Silicon Valley, CA†	41.2%	1	4.64	58.7%	1.3%	$76,290	11.4%
Huntsville, AL	31.2%	0	4.41	45.4%	0.9%	42,035	24.3%
Boulder, CO	42.6%	1	4.30	52.9%	2.0%	53,706	7.8%
Corvallis, OR	34.9%	1	4.28	24.6%	1.3%	48,709	58.1%
Idaho Falls, ID	26.2%	0	4.22	47.5%	0.6%	27,452	12.3%
Binghamton, NY	24.7%	0	3.16	37.4%	0.8%	25,985	9.6%
Wichita, KS	23.2%	0	2.98	33.3%	0.9%	36,773	4.2%
Palm Bay-Melb.-T'ville, FL	28.2%	0	2.97	41.1%	1.0%	28,852	19.6%
Research Triangle, NC‡	32.9%	3	2.68	31.9%	1.2%	96,758	6.4%
Washington, DC-VA-MD-WV	37.7%	2	2.64	40.7%	1.3%	60,757	13.2%
Portland-Vancouver, OR-WA	22.9%	0	1.35	36.3%	1.5%	45,615	16.8%
National metro average	22.5%	< 1	0.78	32.4%	1.0%	32,946	8.5%

Notes: The MSAs are ranked by high-tech location quotient (LQ), creative class and bohemian percentage. § This is the number of Research I-type universities in a given MSA, following the Carnegie Classification of institutions of higher education. † Silicon Valley corresponds to the San Jose-Sunnyvale-Santa Clara, CA MSA.‡ The Research Triangle combines both the Durham, NC MSA and Raleigh-Cary, NC MSA. *Source*: Author's calculations.

It is well known that neither factor endowments nor resulting economic activity are evenly spread across space. Thus in order to understand regional differences in metropolitan economic performance, it is helpful to look at the spatial distribution of the three Ts of development. The MSAs with the highest concentrations in talent, technology and tolerance variables are listed in table 2.1. It is striking that with only three exceptions, these MSAs have enjoyed significant above-average growth from 2000 to 2006.

In order to get a more comprehensive picture of regional growth dynamics, the next section investigates to what extent growth might vary in line with a creative class-based regional typology of metropolitan areas in the U.S.

2.3.2 Spiky regions, but converging?

While the impetus for examining the origins of regional growth is clear, firm agreement on what constitutes a region remains elusive. Since the 1950s the Bureau of Economic Analysis (BEA) has grouped the states into eight regions based primarily on cross-sectional similarities in their socioeconomic characteristics. Recognizing the limitations of this regional classification scheme, several recent studies have

[4] A modified and extended version of this path model and its relevance to the current context are explored in appendix 2.A.

looked to further the understanding of regional composition.

Crone (2005) and Crone and Clayton-Matthews (2005) group states into regions based on the similarities in their business cycles. They apply k-means cluster analysis to the cyclical components of Stock-Watson-type indices estimated at the state level to group the 48 contiguous states into eight regions with similar cycles. Most recently, O'hUallacháin (2008) uses principal components and cluster analyses as a framework for the identification of regions based on state-level growth measures in the U.S.

This paper aims to further narrow the definition of a region by using indicators of technology, talent, and tolerance at the level of the metropolitan statistical area (MSA). The main idea behind a regional typology of high-tech US metropolitan areas is to cluster observations into groups that share similar features and then investigate the way in which the groups differ. Clustering methods are among the most widely used techniques to partition such data into meaningful sub-groups. However, most clustering approaches are largely ad-hoc and lack the rigor usually associated with structural models.

In order to address these issues, I use a procedure that adopts a statistical model to group the data.[5] The model-based clustering procedure is performed with the variables summarised in table 2.2 and leads to the selection of five clusters of MSAs. The variation in the three T-characteristics of these clusters are shown in figure 2.3 below. The "high-tech core" cluster includes well-established high-tech centres such as Silicon Valley, the Research Triangle in North Carolina and the Boston-Cambridge metro area. "High-tech bohemia" metro areas are highly diverse places such a Boise City, ID and Portland, OR, whereas the "industrial core" is made up of old economy regions such as Toledo, OH, Decatur, IL or Scranton, PA. The "creative high-tech" cluster includes established, highly dynamic urban centers including Bolder, CO and Seattle, WA. Upcoming, growing areas with considerable future potential for the development of high-tech like Kansas City, MO make up the "creative periphery".[6]

The great majority of empirical studies that analyse economic growth across countries make use of some version of the so-called convergence equation (e.g. Mankiw, Romer, and Weil (1992) or Sala-i Martin (2006)). In its most general form, this empirical relationship can be written as

[5] Specifically, I follow Fraley and Raftery (2002, 2006) who implement a cluster analysis based on parameterized Gaussian mixture models. Five clusters are obtained with a BIC outcome of -8,325.8 using the VVI parameterization (diagonal, varying volume, varying shape) for the component covariance matrix.

[6] The history of high-tech development in Portland, Boise City and Kansas City, their potential as future leading hubs for the high-tech industry as well as policy options for economic development are discussed in Mayer (2009).

$$\ln y_{it} - \ln y_{it-\tau} = \alpha_i + \beta_i \ln y_{it-\tau} + \gamma X_{it} + \varepsilon_{it} \tag{2.1}$$

where y_{it} is per capita GDP in spatial unit i in period t, X_{it} is a vector of determinants of economic growth that usually includes a time trend and, depending on the underlying model, time-varying accumulation rates which control for differences in the steady states. In line with the literature, I am using the investment share of real GDP as a proxy for the steady state savings rate. α_i is a region- or cluster-specific effect. The summary statistics for the growth model are provided in table 2.2, whereas the estimates for the metropolitan β-convergence are summarised in table 2.3 below.

Table 2.2: Summary statistics – data set used for growth model

Variable	Mean	Std. dev	Min	Max
GDP GROWTH (IN %, 2000-2006)	8.54	7.01	-13.17	45.80
Solow-Swann growth variables				
SAVINGS RATE AS % OF GDP*	22.35	1.54	19.21	25.70
$n+\delta+g^{\dagger}$	6.09	1.25	3.27	15.39
HUMAN CAPITAL‡	18.76	5.86	7.40	43.40
Talent, technology and tolerance				
CREATIVE CLASS (%)	22.52	5.04	0.74	42.88
HIGH-TECH LQ	0.78	0.70	0.05	4.64
BOHEMIANS (%)	0.97	0.34	0.10	2.77
N. Obs.		359		

Notes: * Aggregate gross savings as percent of metropolitan GDP which includes personal savings and those of businesses and government. The MSA savings are proxied by state-level savings rates from the StateMaster Nest Egg Index. † n is the average annual MSA population growth from 2000–2007 and $g+\delta$ are depreciation and investment growth with are assumed to be fixed at 5.00% following Bernanke and Gürkaynak (2001). ‡ In line with the literature, I am using the share of people with a bachelor's degree (above the age of 25 in 2000) as a proxy for the level of human capital. See also the discussion on alternative human capital proxies in footnote 11 in section 2.5.3. *Sources*: Author's calculations using data from BEA and StateMaster.com.

The coefficient of the initial per capita income is negative and significant which is consistent with the standard convergence literature (see e.g. Fousekis (2007), Christopopoulos and Tsionas (2007) or Young, Higgins, and Levy (2008) for recent work on regional convergence in the U.S.). Most interestingly, the coefficient on savings is negative across all specifications, suggesting that – in contrast to the standard predictions of growth theory – the higher a metro area's relative *dissavings*, the higher its rate of GDP convergence.[7] Column (1) shows the simple Solow-Swan benchmark specification of the cross-

[7]This finding is particularly interesting in the context of low household savings, metropolitan housing stress and the origins of

Table 2.3: Estimating β-convergence across metropolitan creative clusters, 2000–2006

	Cross section				
	(1)	(2)	(3)	(4)	(5)
$\ln y_{it-\tau}$	-0.029**	-0.060***	-0.067***	-0.068***	-0.063***
	(0.014)	(0.017)	(0.018)	(0.017)	(0.016)
$\ln(Savings_{it}/GDP_{it})$	-0.191***	-0.212***	-0.254***	-0.267***	-0.212***
	(0.054)	(0.054)	(0.055)	(0.055)	(0.053)
$\ln(n+\delta+g)$	-0.032*	-0.049**	-0.055***	-0.054***	-0.053***
	(0.019)	(0.020)	(0.020)	(0.020)	(0.020)
ln(*Human capital*)*	–	0.045***	0.081**	0.089**	0.029*
		(0.017)	(0.044)	(0.043)	(0.017)
Talent, technology and tolerance variables					
ln(*Creative class*)	–	–	-0.007	-0.011	–
			(0.041)	(0.040)	
ln(*Location quotient*)	–	–	0.015***	0.021***	–
			(0.007)	(0.007)	
ln(*Bohemians*)	–	–	0.026	0.045**	–
			(0.017)	(0.019)	
3T interaction term†	–	–	–	0.011**	–
				(0.005)	
Cluster dummy variables‡					
High-tech bohemia	–	–	–	–	0.035***
					(0.013)
Creative high-tech	–	–	–	–	0.012
					(0.015)
Creative periphery	–	–	–	–	0.014
					(0.010)
High-tech centres	–	–	–	–	0.060***
					(0.023)
Constant	0.883***	1.310***	1.553***	1.663***	1.290***
	(0.195)	(0.249)	(0.242)	(0.244)	(0.222)
Adj. R^2	0.052	0.068	0.113	0.127	0.101
N. Obs.			359		

Notes: * In line with the literature, I am using the population share of college degree holders above the age of 25 in 2000 as proxy of the level of human capital. † The interaction term between the three Ts helps to capture the fact that on their own the Ts might be necessary, but insufficient conditions for growth. A model of such direct and indirect interaction between the three Ts of regional development and regional growth dynamics is presented in appendix 2.A. ‡ The dummy coefficients indicate cluster-specific convergence differences relative to the cluster "industrial core" that contains old economy regions such as Toledo, OH, Decatur, IL or Scranton, PA (see main text for a brief qualitative characterisation of the other clusters). * Significant at the 10% level, ** at the 5% level and *** at the 1% level. *Source*: Author's calculations using data from BEA and StateMaster.com.

sectional convergence equation imposing full parameter homogeneity, whereas column (2) introduces human capital and column (3) adds the three Ts of economic development, i.e. the creative class as measure for talent, the high-tech location quotient (LQ) as measure for technology, and the share of bohemians as proxy for tolerance. The model in column (4) incorporates an interaction term between the three Ts which helps to capture the fact that on their own the Ts might be necessary, but insufficient conditions for growth.[8] Models (3) and (4) confirm that technology is a key driver for regional GDP growth and that the simultaneous interaction between talent (creative class), technology and tolerance also plays an important role in this process. Indeed, the positive and significant coefficient on the interaction term suggests that there are important complementarities with regard to the interplay between the three Ts of regional development and regional growth dynamics.[9]

At the same time however – in line with the evidence of Florida, Mellander, and Stolarick (2008) who find that human capital and the creative class affect regional development through different channels – the creative class measure on its own does not influence the regional convergence of GDP. Since conventional human capital measures perform better in accounting for regional income and the creative class is more suited with regard to accounting for regional labor productivity, I am using the creativity-based clusters to allow for the presence of both channels in column (5) without introducing econometric specification issues.

The significance of the cluster dummy coefficients is mixed which indicates that a given MSA's speed of convergence depends on its cluster membership. Most importantly, there is evidence that both the "high-tech core" and "high-tech bohemia" exhibit per capital growth convergence rates that are significantly above those of the MSAs in the "industrial core". In other words, the landscape of the metropolitan U.S. is characterised by two secular trends. First, there is the within-convergence of GDP among the centres in the "high-tech core" and those in the "high-tech bohemia" cluster, which include Silicon Valley, the Research Triangle and the Boston High-Tech in the former and Boise City, Colorado Springs, and Portland-Vancouver-Beaverton in the latter. Second, the other clusters show lower speeds of convergence, supporting the notion of spiky economic development.

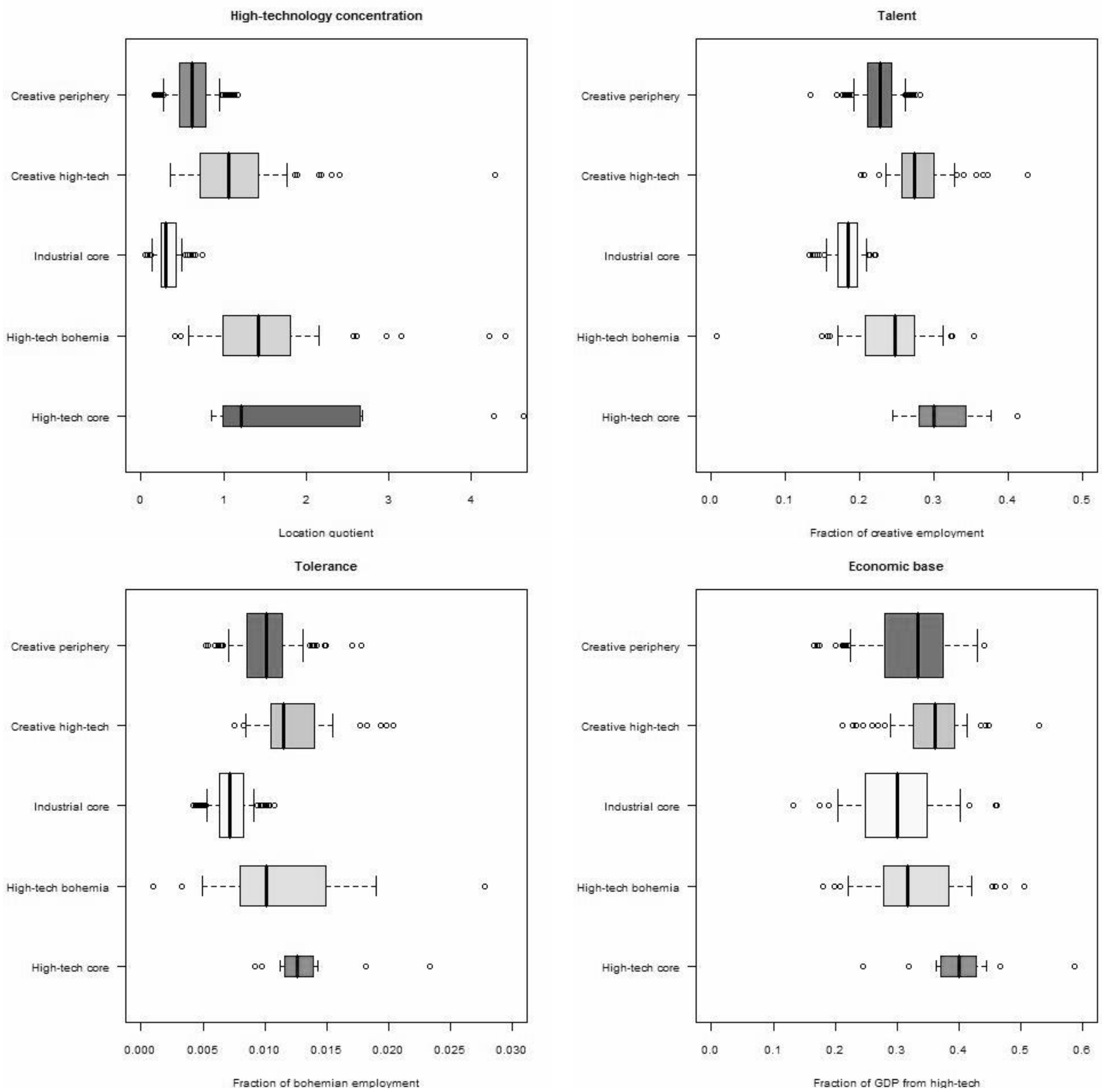
On the surface, this seems to run against the grain of received wisdom of "old geography" thinking

the current financial crisis. However, further exploration is beyond the scope of this paper and will be pursued elsewhere.

[8] I am grateful to one of the reviewers for suggesting this specification.

[9] In fact, this is approach is related to, albeit less complex than the structural equation model estimated in Florida, Mellander, and Stolarick (2008). A modified version of their model – augmented by quality-of-life variables and GDP growth – is also estimated for exploratory purposes. The results are presented in appendix 2.A

Figure 2.3: Regional variation of talent, technology and tolerance



Notes: The boxplots show the distribution of key variables across the five types of US metropolitan high-tech industry clusters. The clusters were identified using a model-based cluster analysis. *Source*: Author's calculations.

which views globalisation as a "flattener" in the sense of Friedman (2005). Thus instead of bringing about a more uniform distribution of economic development across space, the world of specialised high-tech activity is certainly becoming more spiky or curved, rather than flat – a view long held by Florida (2005) and recently refined by McCann (2008). Indeed, using regional estimates of economic activity, Florida, Gulden, and Mellander (2008) identify 12 mega-regions for North America, i.e. regions that not only show large agglomerations of people, but also have large markets, substantial innovation and highly-skilled talent.

Shifting focus from MSA-level specialisation and concentration to county-level data on individual firms, the next section increases the level of both spatial and economic resolution in the investigation of the link between regional growth and the three Ts of economic development. In particular, I am now testing how the location of new high-technology plant openings is related to university R&D and the creative class.

2.4 Firm location as discrete choices

There is a significant amount of empirical research investigating firms' location choices. Typically, the most appealing model for discrete data is the conditional logit model (CLM). Carlton (1983) is one of the first to investigate the problem of firm location using the discrete choice Random Utility Model (RUM) framework pioneered by McFadden (1974). The popularity of this approach stems from the fact that there is a direct link between the framework of random utility (profit) maxismisation and its econometric specification.

2.4.1 The conditional logit model

Considering J spatial choices with $j = 1, \ldots, J$ and N firms with $i = 1, \ldots, N$, then the profit derived by firm i in sector k if it locates in county j is given by

$$\pi_{ijk} = \beta' \boldsymbol{z}_{ij} + \varepsilon_{ijk}, \tag{2.2}$$

where $\boldsymbol{z}_{ij}$ is a vector of explanatory variables specific to each county (and possibly sector), β is a vector of unknown coefficients and ε_{ijk} is an i.i.d. stochastic term assumed to have an Type I extreme values distribution. Then, following McFadden, the probability of any one firm from sector k locating in county j is

$$p_{ijk} = \frac{\exp(\beta' \boldsymbol{z}_{jk})}{\sum_{j=1}^{J} \exp(\beta' \boldsymbol{z}_{jk})}. \tag{2.3}$$

Then, letting $y_{ij} = 1$ if a firm locates in county j and $y_{ij} = 0$ otherwise, the log likelihood of the CLM then takes the familiar form of $LL_{CLM} = \sum_i^N \sum_j^J y_{ij} \log p_{ij}$

However, one of the restrictions of the CLM is the assumption that the location decisions of firms are independent, once all relevant factors are accounted for. This implies that the spatial distribution of new firm formation is determined by a multinomial law. The high level of clustering presented in much of the literature on firm location choice (including the evidence contained in the first part of this paper) suggests that this assumption might be too restrictive, since the regressors might not be able to account for the clustering.

Indeed, theory suggests that much of this clustering might be due to (unobserved) regional heterogeneity due to localisation economies or other specific local factors. As a consequence, the data might exhibit variances larger than those permitted by the binomial model. This results in *overdispersion* and means that the standard errors of the CLM will be underestimated.[10] In an extension to the CLM, the Dirichlet-Multinomial model (DM) permits for overdispersed data by including a random effect α_{jk} in equation 2.2.

Ignoring this random effect for the time being, the CLM assumes that firms' decisions are based on choice-specific attributes common to all firms irrespective of their industry. In this case, $\boldsymbol{z}_{ij} = \boldsymbol{z}_j$ and so the log-likelihood function of the CLM can then be expressed as

$$\begin{aligned} LL_{CLM} &= \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij} \log p_{ij} = \sum_{j=1}^{J} n_j \log p_j \\ &= \sum_{j=1}^{J} n_j (\beta' \boldsymbol{z}_j) - \sum_{i=1}^{N} n_j \log \left[\sum_{j=1}^{J} \exp(\beta' \boldsymbol{z}_{jk}) \right], \end{aligned}$$

[10]While overdispersion has consequences similar to those of the presence of heteroskedasticity in linear regression models, the CLM estimator is still consistent if the conditional mean, i.e. $\exp(\beta' \boldsymbol{z}_{jk})$, is correctly specified.

where n_j is the number of new firm births in county j.

2.4.2 The relation between the conditional logit and the Poisson regression

In the context of firm location choice, large choice sets are not uncommon and it may be very cumbersome to estimate a conditional logit model. In the WFG data, for example, over twelve thousand new firm investments (see table 2.5) across more than three thousand U.S. counties would yield a $N \times J$ choice set in excess of *36 million observations*. Similar problems of infeasibly large choice sets are discussed in Ellickson, Houghton, and Timmins (2010).

However, Guimarães, Figueiredo, and Woodward (2003, GFW hereafter) demonstrate that, by taking advantage of an equivalence between the log-likelihood of the CLM and the Poisson regression for count data[11], the CLM can be estimated regardless of the number of choices. Following this logic, the key link between the CLM and Poisson regressions can be outlined as follows.

Let the number of firm births in a given county n_j be independently Poisson-distributed with $E(n_j) = \lambda_j = \exp(\beta' \boldsymbol{z}_j)$. Then, the log-likelihood function of the Poisson regression model can be written as

$$
\begin{aligned}
LL_P &= \sum_{j=1}^{J} (-\lambda_j + n_j \lambda_j - \log n_j!) \\
&= \sum_{j=1}^{J} [-\exp(\beta' \boldsymbol{z}_j) + n_j \exp(\beta' \boldsymbol{z}_j) - \log n_j!]. \qquad (2.4)
\end{aligned}
$$

From first-order conditions it then follows that – with the exception of constant terms – the log-likelihood function of CLM is equivalent to the log-likelihood function of the Poisson distribution.

Consequently, estimates obtained for β are the same in both models. GFW show that the estimated covariance matrix will also be identical in both models under certain conditions. Furthermore, GFW also establish analytical equivalence between the CLM and Poisson models in the more complex approach in which each locational decision is based on choice-specific attributes common to groups of firms, i.e. $\boldsymbol{z}_{ij} = \boldsymbol{z}_{jk}$

Using this key insight then provides an empirical strategy for estimating the CLM in the context of large choice set such as in WFG.

[11] Hellerstein and Mendelsohn (1993) provide a general description of count data models for microeconometric applications.

2.4.3 Overdispersed count data

As indicated above, the potential problem of overdispersion in spatially diverse count data imposes limitations on the statistical inference that can be based on the CLM. The issue of overdispersion can be addressed by assuming that the n_{jk} firm births follow a Dirichlet-Multinomial distribution instead of a multinomial distribution as implied by the CLM. Using a similar approach to GFW, Guimarães and Lindrooth (2007) provide the analytical equivalence between the log-likelihood of the Poisson distribution, the log-likelihood of the CLM conditional on random effects, and the unconditional log-likelihood of the DM. Since the Dirichlet-Multinomial can be shown to nest the CLM, a straightforward log-likelihood ratio test provides the direct means of testing for overdispersion.

Examining the role and potential for replication in economics, Hamermesh (2007) points out the paucity of both pure replication (checking published papers using their data) and scientific replication (using data representing different populations). This warrants a brief account of my own replication strategy. First, I coded the CLM in Matlab in order to replicate the CLM results reported by WFG. Then, to account for overdispersion, I used the original code of WFG to estimate the DM in Stata in order to analyse the joint impact of R&D and the creative class.

2.5 R&D, spillovers and firm-location choice

I obtained the full data set used in WFG which contains individual establishment data by county and a set of establishment characteristics that may affect the firm's profit function. In addition to the county characteristics contained in WFG, I am adding a county-level measure of Florida-type creativity.

2.5.1 Data overview

The full WFG dataset is relatively large and contains over 60,000 observations of manufacturing firms for the 3,067 counties belonging to the 48 contiguous US states. The main variables used in the WFG dataset are listed below.[12]

HTBIRTH Dependent variable, count of firm formations,

[12]The summary statistics of the full data set and a detailed description of the construction of the WFG variables are presented in the appendix of an earlier version of this paper.

LABOUR COST	County wage and salary earning per job in 1996,
LAND COST	County population density in 1996,
TAXES	County property taxes in 1997,
URBANISATION	County density of manufacturing and service establishments per km^2 in 1996,
LOCALISATION	Number of establishments in same 2-digit SIC industry,
WTPI	Weighted market size $WTPI_c = \sum_j^J TPI_j / d_{cj}^{\tau}$, where TPI_j is total personal income for county j and d_{cj} is the linear distance between the centroids of county c and j. τ is a distance decay parameter set equal to 2,
AMENITIES	County natural amenities,
HUMAN CAPITAL	College graduates as percentage of population, (25 years +) in 1990,
UNI R&D	Weighted university R&D expenditure, $WRD_c = \sum_j^J \delta_{cj} RD_j$.

As highlighted above, my main hypothesis is to test if – in addition to R&D expenditure – the presence of "creative class" human capital – plays a role in new high-tech firm formation. Thus, I am including an additional regressor to the relevant factors of location specific characteristics.

CREATIVE CLASS	Share of creative class employment per county, based on USDA-ERS methodology.

After dropping all non-high-tech observations, the data used in the estimations consists of 21,469 observations and is summarised in table 2.4.

2.5.2 Dependent Variable

For the dependent variable – the number of manufacturing high-tech firm births – WFG use establishment level (3-digit SIC) data from the U.S. Census Bureau, encompassing all new known entrants in each US county from 1997 to 2000. High-tech firm births are identified using the BLS's classification of high-technology industries, which is based on measures of employees engaged in R&D activities.

The WFG dataset contains a total number of 12,007 new high-tech plant openings in the manufacturing sector. Table 2.5 lists these high-tech firm births across different industry groups. Indeed, figure 2.4 shows the spatial distribution for new high-tech plants and their relationship with universities' R&D spending. MSAs with the lowest R&D expenditures by academic institutions are shaded in dark grey, those with the highest expenditures are shaded in light grey. As can be seen in this figure, new firm

Table 2.4: Summary statistics – data set used for firm entry model

Variable	Mean	Std. dev	Min	Max
HIGH-TECH BIRTHS	0.56	3.32	0.00	214.14
TOTAL UNIVERSITY R&D	5,669.24	38,349.85	0.00	77,3596.40
WEIGHTED UNIVERSITY R&D	73,440.66	163,318.50	0.00	1,219,169.69
LOG WU R&D	-22.25	29.55	-48.35	14.01
Local cost factors				
LABOUR COST	9.94	0.20	9.32	10.91
LAND COST	2.74	1.63	-2.51	9.95
TAXES	6.30	0.65	3.39	9.42
Local demand				
WEIGHTED MARKET SIZE	20.16	1.44	16.08	26.14
NATURAL AMENITIES	1.21	0.30	0.00	1.95
Agglomeration factors				
LOCALISATION ECONOMIES	-28.53	24.80	-59.11	0.03
URBANISATION ECONOMIES	-1.68	1.71	-7.77	6.96
*Talent factors**				
HUMAN CAPITAL (HS)†	69.53	10.35	31.60	95.50
HUMAN CAPITAL (BA)‡	13.39	6.44	3.70	53.40
CREATIVE CLASS	14.32	5.04	1.75	42.88
N. Obs.		21,469		

Notes: All variables are in logs, unless otherwise stated such as where * indicates a percentage (specifically, the % of total labour force or population). † Human capital defined in terms of high-school graduates as a percentage of the population (25 year +) in 1990. ‡ Human capital defined defined in terms of college graduates with a bachelor's degree as a percentage of the population (25 year +) in 1990. *Sources*: Woodward, Figueiredo, and Guimarães (2006), USDA-ERS, author's calculations.

Table 2.5: High-tech plant openings by 2-digit SIC industries (1997–2000)

SIC	Industry	Number of plants	(%)
28	Chemicals	2,367	19.7
29	Petroleum refining	104	0.9
34	Ordnance and accessories	122	1.0
35	Industrial machinery and equipment	3,020	25.1
36	Electronic & other electrical equipment	2,815	23.5
37	Transportation equipment	1,544	12.8
38	Instruments and related products	2,035	17.0
Total all R&D intensive industries		12,007	100.0

Sources: Author's calculations using data from Woodward, Figueiredo, and Guimarães (2006).

formations are highly concentrated around some MSAs – particularly those with high R&D expenditures – suggesting the presence of localised spillovers.

Figure 2.4: High-technology plant openings by MSA (1997–2000)

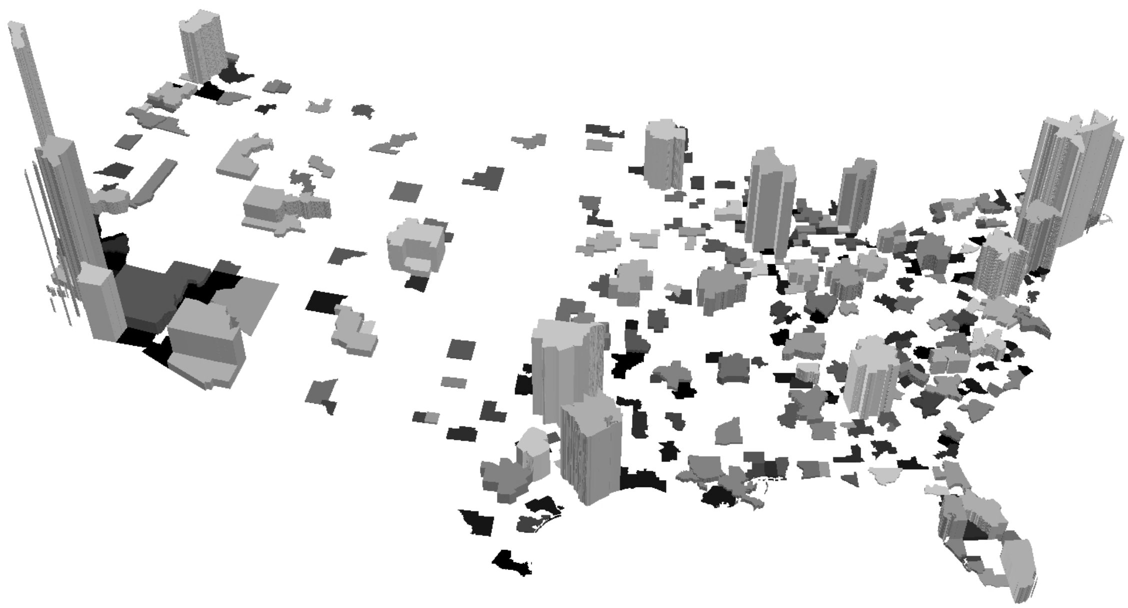
Notes: The height of the columns corresponds to the number of new high-tech firm births. MSAs ranked in the top quintile with regard to university R&D expenditures are in light grey, the lowest quintile is in dark grey. *Source*: Author's calculations.

2.5.3 Independent variables

Creative class index

The popularity of Florida's creative class measure has given rise to a number of empirical applications of this concept. Most recently, the Economic Research Service of the U.S. Department of Agriculture (USDA-ERS) has published a county-based creative class measure, which excludes from the original Florida measure many occupations with low creativity requirements and those involved primarily in economic reproduction (i.e., numbers proportional to population). Figure 2.5 illustrates the spatial distribution by county for the creative class measure. A first visual comparison with figure 2.4 suggests similar spatial patterns for new firm formations, industry concentration and creativity, the key assertion made by Lee, Florida, and Acs (2004). Beyond these spatial patterns, firm births are likely to depend on the presence of agglomeration effects from industry localization. Such localization externalities are clearly made visible in figure 2.5 by scaling the number of new firm formations by the high-tech industry location quotient.

Figure 2.5: High-tech plant openings, industry concentration and the creative class MSAs.



Notes: The height of the columns corresponds to the number of new high-tech firm births scaled by the high-tech industry location quotient. MSAs ranked in the top quintile with regard to the creative class measure are in light grey, the lowest quintile is in dark grey. *Source*: Author's calculations, USDA-ERS using data from the U.S. Census Bureau.

Using the additional variables from the cluster analysis, I re-examine the hypothesis set out in Florida (2002a) with the WFG dataset and test if – in addition to university R&D – the presence of "creative class" positively affects firm decisions to locate new establishments in high-technology industries, after controlling for other important county and regional cluster characteristics.

Florida's original data has the major drawback that it excludes rural areas as it is aggregated at the level of Metropolitan Statistical Areas instead of counties. For the purpose of this paper, I will thus rely on the data published by the USDA-ERS which uses data from the U.S. Census Bureau. This choice is also supported by recent evidence in Wojan, Lambert, and McGranahan (2007) that this new measure conforms more closely to the concept of creative class and proves to be more highly associated with regional development than the original Florida measure.

Other independent variables

One of the key econometric challenges is to distinguish between different types of unobserved spillover effects, particularly since University R&D is likely to be correlated with the presence of a bohemian effect. Thus in order to avoid any endogeneity bias due to omitted variables, both R&D and the creativity measures are included in the estimation.

Table 2.6: Correlations of "milieu" variables

Variable		Creative Class	Localisation	Urbanisation	Amenities	Human Capital* HS	BA	University R&D
Creative Class		1.00	–	–	–	–	–	–
Localisation		0.38	1.00	–	–	–	–	–
Urbanisation		0.64	0.53	1.00	–	–	–	–
Amenities		0.18	-0.01	-0.07	1.00	–	–	–
Human Capital	HS	0.62	0.22	0.24	-0.04	1.00	–	–
	BA	0.88	0.27	0.45	0.15	0.69	1.00	–
University R&D		0.31	0.26	0.46	-0.06	0.12	0.21	1.00

Notes: * The population percentage of high-school graduates HS provides a broader alternative proxy measure for human capital compared to the percentage of university graduates BA. Both measures are tested and the results are stable for both specifications of human capital. *Sources*: Author's calculations using data from Woodward, Figueiredo, and Guimarães (2006) and USDA-ERS.

Furthermore, the creativity measure is also likely to be correlated with some of the other milieu-based variable used in WFG. Indeed, as can be seen from table 2.6, perhaps unsurprisingly Creative Class is most highly correlated with Urbanisation and Human Capital[13], a reflection of the well-known urban and high-human capital bias of Florida's argument. While there is correlation with R&D, this correlation seems in line with the other independent variables used by WFG.

[13]WFG use a very broad measure of human capital which is defined as the population percentage of high-school graduates. In line with the majority of the literature, however, I define human capital as the share of university degree holders for the remainder of the analysis. Both measures are tested and the results are largely stable for both specifications of human capital. See table 2.4 for the summary statistics of the two measures.

2.6 Firm-level evidence

Table 2.7 presents the estimates for the location determinants of high-tech firm births in US counties, using both the DM and CLM regressions. While the likelihood ratio test provides evidence for overdispersion, the parameter estimates vary only slightly between both models. With the exception of the control for the cost of labour, all estimates are significant across all estimations.

2.6.1 General results

Column (1) repeats the CLM results from WFG, whereas my replication with the share university degree holders as human capital proxy instead of high-school graduates is presented in column (2). The estimation results for the model that includes the creative class measure are shown in column (3). Columns (4) and (5) show the results using the DM which accounts for overdispersion. The new human capital proxy yields results broadly similar to those in WFG, providing supporting evidence that both location factors and cost factors have a significant impact on the number of new high-tech firm openings. The estimated coefficient for the spatially weighted market size also proves to be statistically significant and has the expected sign. There is also evidence that the two agglomeration variables – LOCALISATION and URBANISATION – affect high-tech investors locational decisions. Furthermore, the presence of qualified labour and natural amenities also have an influence on new firm formation. UNIVERSITY R&D, the key variable of interest in WFG, is positive and statistically significant, suggesting that a closely situated university plays a important role in explaining high-tech location decisions.

The main variable of interest in this paper, CREATIVE CLASS, is positive and statistically significant in both the CLM and the DM specification. Adding this variable does not significantly change the estimated coefficients with regard to the variables in the WFG model. These general results seem to provide considerable support to the original hypothesis that the presence of "creativity and ideas" may indeed generate new economic activity at the firm level. Most importantly, perhaps, the null hypothesis that there is no relation can be rejected. Since the coefficients in the CLM and DM have a direct economic interpretation, the results suggest that, ceteris paribus, a 10 percent increase in a given county's share of creative employment leads to an increase in the probability of location in this county that ranges between 4.07 and 5.34 percent.

Table 2.7: Location determinants of U.S. high-tech plant openings – all sectors

	WFG06	*Replication*		*Replication*	
Model	CLM	CLM	CLM	DM	DM
Variables	(1)	(2)	(3)	(4)	(5)
LABOUR COST	-0.0010*	0.0399*	0.0144*	-0.2805	-0.3033
	(0.077)	(0.078)	(0.077)	(0.099)	(0.101)
LAND COST	-0.4589	-0.5180	-0.4981	-0.5519	-0.5319
	(0.047)	(0.490)	(0.047)	(0.059)	(0.060)
TAXES	-0.1380	-0.1030	-0.1122	-0.0706	-0.0717
	(0.024)	(0.023)	(0.023)	(0.028)	(0.028)
WEIGHTED MARKET SIZE	1.0885	1.0951	1.0779	1.0857	1.0743
	(0.013)	(0.013)	(0.013)	(0.016)	(0.016)
LOCALISATION ECONOMIES	0.0215	0.0225	0.02147	0.0190	0.0185
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
URBANISATION ECONOMIES	0.2663	0.3005	0.2827	0.3536	0.3330
	(0.042)	(0.044)	(0.045)	(0.054)	(0.055)
HUMAN CAPITAL†	0.0098	0.0036	0.0028	-0.0021*	-0.0121
	(0.002)	(0.002)	(0.001)	(0.002)	(0.004)
NATURAL AMENITIES	0.2435	0.2350	0.2096	0.1304	0.1071
	(0.032)	(0.032)	(0.033)	(0.040)	(0.041)
WEIGHTED UNIVERSITY R&D	0.0032	0.0036	0.0034	0.0029	0.0029
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
CREATIVE CLASS	–	–	0.5335	–	0.4071
			(0.097)		(0.115)
Log-likelihood	-11,522.5	-11,540.9	-11,525.9	-10,936.2	-10,929.3
LR overdispersion test χ^2				994.4	1,093.6
Radius in miles (δ)				60	60
N.Obs.			21,469		

Notes: Estimated models are the conditional logit model (CLM) and the Dirichlet-Multinomial model (DM). Standard errors in parentheses. † Human capital is proxied by the share of university degree holders in the specification of columns (2)–(5), whereas the original WFG specification uses the share of high-school graduates. All estimates are significant at the 1% level, except those marked by *. *Sources*: Author's calculations.

While a similar interpretation is not possible for the R&D coefficient due to its weighted nature[14], WFG report simulation results which reveal that a $1 million increase in university spending on R&D increases the probability of location by less than 0.1 percent.

Overall, the results from including a measure of creativity are encouraging. In addition to the presence of university R&D, I find evidence for significant creativity-based localised spillovers that have an impact on the investment decision on high-tech firms.

2.6.2 Industry-level results

Table 2.8: Location determinants of U.S. high-tech plant openings – industry results

	Conditional logit model – Poisson regression estimates				
Variables	SIC 28$^\flat$	SIC 35$^\natural$	SIC 36$^\sharp$	SIC 37†	SIC 38‡
LABOUR COST	-0.3354*	-0.0565	1.0494***	-0.6983***	-0.8611***
	(0.174)	(0.162)	(0.171)	(0.220)	(0.191)
LAND COST	-0.5963***	-0.4636***	-0.2639**	-0.9888***	-0.2512***
	(0.112)	(0.099)	(0.106)	(0.138)	(0.121)
TAXES	-0.0491	-0.1092**	-0.1175**	-0.3233***	0.0862
	(0.052)	(0.046)	(0.051)	(0.062)	(0.059)
WEIGHTED MARKET SIZE	1.0921***	1.0819***	1.0576***	1.1502***	1.0991***
	(0.029)	(0.026)	(0.026)	(0.037)	(0.032)
LOCALISATION ECONOMIES	0.0170***	0.0361***	0.0283***	0.0181***	0.0153***
	(0.002)	(0.008)	(0.004)	(0.003)	(0.003)
URBANISATION ECONOMIES	0.4554***	0.2722***	-0.0185	0.7218***	0.0530
	(0.102)	(0.091)	(0.096)	(0.129)	(0.110)
HUMAN CAPITAL	-0.0136**	-0.0260***	-0.0071	-0.0471***	0.0108
	(0.007)	(0.006)	(0.007)	(0.009)	(0.007)
NATURAL AMENITIES	0.0884	-0.0408	0.6272***	-0.1019	0.3459***
	(0.073)	(0.065)	(0.069)	(0.093)	(0.080)
WEIGHTED UNIVERSITY R&D	0.0010	0.0023**	0.0053***	0.0059***	0.0040***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
CREATIVE CLASS	0.0692	0.8433***	1.2875***	0.4281**	0.8630***
	(0.211)	(0.190)	(0.233)	(0.206)	(0.247)
Log-likelihood	-2,117.6	-2,480.0	-2,072.2	-1,995.4	-1,705.6

Notes: ♭ Chemicals, ♮ industrial machinery and equipment, ♯ electronic & other electrical equipment, † transportation equipment and ‡ instruments and related products. Standard errors in parentheses. * Significant at the 10% level, ** at the 5% level and *** at the 1% level. *Source*: Author's calculations.

In order to determine how these results change according to industry, I estimate the CLM model sep-

[14] In table 2.7, the value of δ radius is set equal to 60 miles, the value that maximises the log-likelihood function.

arately for two-digit SIC industries where there is sufficient data within each SIC category.[15] As for the overall results, most of the coefficient estimates have the expected signs, although their relative importance varies significantly across industries. The results are presented in table 2.8.

In contrast to the overall results, not all "milieu variables" enter significantly across industries. While URBANISATION ECONOMIES and HUMAN CAPITAL show an alternating pattern of significance in the different industries, CREATIVE CLASS is significant in all but the chemical industry (SIC 28). The significance of university R&D is similar to the industry-level results reported in WFG.[16] Apart from industry-specific differences, one possible alternative explanation is certainly the high levels of correlation among those regressors. This is something that would need further inquiry. Indeed, this suggests that, in addition to using separate measures for urbanisation and the level of human-capital, further insights can be gained in to the firm location decision process by adding a measure that proxies for creativity. Initial explorations for the industry-level data indicate similar results from the DM model, thus suggesting that overdispersion may not be the key driver of these results.

2.7 Conclusion and outlook

The key hypothesis tested in this paper links the presence of localised spillovers from "creative people" to different measures of economic performance in the knowledge-intensive, high-tech sector. These effects appear to be robust at different levels of aggregation and across measures, namely for urban GDP at the MSA-level and for new firm formations. In contrast to other theories of agglomeration and regional specialisation, the evidence presented in this paper suggests that the creative class theory is independent of the spatial level of aggregation.[17] In the more specific context of universities as locational anchors for growth and their interaction with talent and technology, my results are able to confirm the importance of university-based R&D for the location choices of knowledge-based firms. At the same time, however, I also find evidence for the creative class conjecture that underlines the importance of a "bohemian milieu" for economic development.

[15] In line with WFG, I excluded SIC 29 and SIC 34 from the industry analysis, because of the small number of plant openings in these sectors (see table 2.5).

[16] I am holding δ constant at the overall optimal level of 60 miles, instead of recomputing optimal levels for all industries as is done in WFG.

[17] See Brakman, Garretsen, and van Marrewijk (2009) for a discussion of how theories that explain the uneven spatial distribution of economic activity (mainly urban economics and New Economic Geography) differ with regard to the relevance they assign to spatial linkages across various levels of spatial aggregation.

At this stage, a number of qualifications are important, particularly with a view for avenues of future research. The firm's location choice problem seems highly endogenous with agglomeration effects, bohemian effects and R&D effects which are probably all interrelated. In order to identify these effects separately, further analysis in the context of a structural equilibrium model of firm sorting could provide further insights. The work of Timmins (2005) provides a starting point in this direction. Bayer and Timmins (2005, 2007) develop the equilibrium properties of such sorting models that allow a firm's location decision to depend on both fixed local attributes (including unobserved attributes) and local interactions and provide a test for uniqueness in empirical analyses of sorting equilibrium. Similarly, Koo (2007) presents a simultaneous equations framework that takes into account the endogeneity of spillovers and agglomeration. He finds that diversity and specialization are important, but the magnitude of their importance decreases as the knowledge intensity of an industry increases.

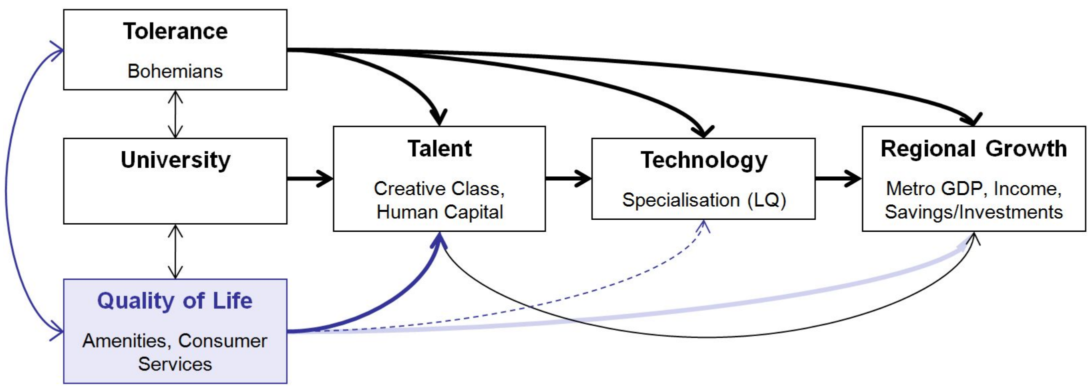
In addition, there is also the issue of measurement of knowledge-based spillovers. In this context, the knowledge capital model due to Griliches (1979) provides the cornerstone of empirical work that investigates the relationship between R&D activity and productivity. There are two distinct notions of R&D spillovers which are often confused in the literature. The first one, R&D intensive inputs purchased from other firms at less than their full quality price, is really a measurement issue (i.e. lack of hedonic price indices that reflect improvements in quality), rather than a true spillover. True spillovers, says Griliches, are "the ideas borrowed by the research teams of industry i from the research results of industry j" (Griliches, 1979, p.104). In a similar vein, Scherer (1982) provides empirical evidence for the significant specification errors that arise if these separate effects are not disaggregated.

And finally, it would be interesting to examine the stability of the firm-level microdata results among a different universe of high-tech firms, in particular among the emerging group of *non-manufacturing high-tech firms*. As the importance of non-manufacturing high-tech increases in the knowledge economy of the 21st century, a closer look at these industries might yield additional valuable findings for policy makers.

2.A Channels of regional development

The direct and indirect interaction between the three Ts of regional development and regional growth dynamics can be further analysied in the path model framework proposed by Florida, Mellander, and Stolarick (2008). A modified and extended version of their model of a regional development system – augmented by quality-of-life variables and GDP growth – is illustrated in figure 2.6.

Figure 2.6: Path model of regional growth.



Notes: This stylised model illustrates the interaction between different regional variables and growth-related economic outcomes. The strength of the arrows indicates the relative importance of the different paths on the basis of the tentative results in table 2.9. *Source*: Author's illustration adapted from Florida, Mellander, and Stolarick (2008).

Its empirical specification can be written as:

$$\log \text{TALENT}_i = \beta_1 \log \text{TOLERANCE}_i + \gamma_1 \log \text{UNI}_i + \delta_1 \log \text{QUALI}_i + \varepsilon_1 \quad \text{(A-1)}$$

$$\log \text{TECH}_i = \beta_2 \log \text{TOLERANCE}_i + \gamma_2 \log \text{TALENT}_i + \delta_2 \text{QUALI}_i + \varepsilon_2 \quad \text{(A-2)}$$

$$\log \text{GROWTH}_i = \beta_3 \log \text{TOLERANCE}_i + \gamma_3 \log \text{TALENT}_i + \delta_3 \text{TECH}_i + \lambda_1 \text{MACRO}_i + \varepsilon_3, \quad \text{(A-3)}$$

where

TALENT	Creative class variable in 2000,
TECH	Technological specialisation, measured as LQ of high-tech industry,

TOLERANCE	Share of Bohemians in 2000,
GROWTH	Metropolitan GDP growth from 2000–2006,
UNI	Number of research I-type universities,
QUALI	Quality-of-Life variables, including the number of restaurants, museums, local taxes, environmental and natural amenities following Blomquist, Berger, and Hoehn (1988) and Shapiro (2006),
MACRO	Macroeconomic variables such a the savings rate, investment growth and population growth.

As a first approach which incorporates only limited interdependencies, this system is estimated using seemingly unrelated regression (SUR) and the results are shown in table 2.9 below. These estimates form the basis for tentative inference on the relative importance of the different paths of regional development, as highlighted by the strength of the arrows in figure 2.6. More comprehensive modelling in a dynamic framework is left for future research.

2.B Acknowledgements

I would like to thank the guest editors, Kevin Stolarick, Charlotta Mellander and Richard Florida, and two anonymous referees for their helpful comments. I am grateful to Nick Kuminoff and Heike Mayer for insights, discussion and vital suggestions on an earlier draft. Inputs from participants at the 2008 Southern Regional Science Association Conference and the 2008 Annual Meetings of the Mid-Continental Regional Science Association are also acknowledged. I am indebted to Paulo Guimarães for sharing the data on new firm formations.

2.C References

ACS, Z. J., AND D. B. AUDRETSCH (1988): "Innovation in Large and Small Firms: An Empirical Analysis," *American Economic Review*, 78(4), 678–690.

ACS, Z. J., D. B. AUDRETSCH, AND M. P. FELDMAN (1992): "Real Effects of Academic Research: Comment," *American Economic Review*, 82(1), 363–367.

ANDERSSON, M., AND K. HELLERSTEDT (2009): "Location Attributes and Start-ups in Knowledge-Intensive Business Services," *Industry and Innovation*, 16(1), 103–121.

Table 2.9: Path analysis for thee Ts and regional growth

Variables	Talent CREATIVE CLASS	Technology LQ	Regional Growth GDP GROWTH
CREATIVE CLASS	–	3.062***	-0.007
		(0.238)	(0.040)
LQ	–	–	0.073***
			(0.023)
BOHEMIANS	0.313***	-0.456***	0.038**
	(0.026)	(0.140)	(0.018)
UNIVERSITY	-0.055***	–	–
	(0.017)		
RESTAURANTS	-0.052***	-0.007	–
	(0.016)	(0.072)	
THEATRES	0.083***	0.113*	–
	(0.014)	(0.068)	
LOCAL TAXES	0.143***	-0.197	–
	(0.024)	(0.114)	
COAST	-0.040***	0.086	–
	(0.013)	(0.060)	
NAT. AMENITIES	-0.047**	-0.039	–
	(0.019)	(0.086)	
TYPOGRAPHY	0.004	0.086***	–
	(0.007)	(0.031)	
GDP LEVEL	–	–	-0.068***
			(0.017)
SAVINGS	–	–	-0.268***
			(0.054)
$\ln(n+\delta+g)$	–	–	-0.054***
			(0.019)
HUMAN CAPITAL	–	–	0.095**
			(0.043)
3T INTERACTION	–	–	Y
Constant	-1.326***	3.451**	1.666***
	(0.286)	(1.341)	(0.241)
RMSE	0.108	0.494	0.064
(Pseudo) R^2	0.733	0.574	0.146
N.Obs.		359	

Notes: Standard errors in parenthesis. * Significant at the 10% level, ** at the 5% level and *** at the 1% level. *Source*: Author's calculations.

ANSELIN, L., A. VARGA, AND Z. ACS (1997): "Local Geographic Spillovers between University Research and High Technology Innovations," *Journal of Urban Economics*, 42(3), 422–448.

——— (2000): "Geographical Spillovers and University Research: A Spatial Econometric Analysis," *Growth and Change*, 31(4), 501–515.

AUDRETSCH, D. B., AND M. P. FELDMAN (1996): "R&D Spillovers and the Geography of Innovation and Production," *American Economic Review*, 86(3), 630–640.

BANIA, N., R. W. EBERTS, AND M. S. FOGARTY (1993): "Universities and the Startup of New Companies: Can We Generalize from Route 128 and Silicon Valley?," *Review of Economics and Statistics*, 75(4), 761–766.

BAYER, P., AND C. TIMMINS (2005): "On the Equilibrium Properties of Locational Sorting Models," *Journal of Urban Economics*, 57(3), 462–477.

——— (2007): "Estimating Equilibrium Models of Sorting Across Locations," *Economic Journal*, 117(2), 353–374.

BEAUDRY, C., AND A. SCHIFFAUEROVA (2009): "Who's Right, Marshall or Jacobs? The Localization Versus Urbanization Debate," *Research Policy*, 38(2), 318–337.

BERNANKE, B. S., AND R. S. GÜRKAYNAK (2001): "Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously," *NBER Macroeconomics Annual*, 16, 11–57.

BLOMQUIST, G. C., M. C. BERGER, AND J. P. HOEHN (1988): "New Estimates of Quality of Life in Urban Areas," *American Economic Review*, 78(1), 89–107.

BLUMENTHAL, P., H. L. WOLMAN, AND E. HILL (2009): "Understanding the Economic Performance of Metropolitan Areas in the United States," *Urban Studies*, 46(1), 605–627.

BOSCHMA, R. A., AND M. FRITSCH (2009): "Creative Class and Regional Growth: Empirical Evidence from Seven European Countries," *Economic Geography*, 85(4), 391–423.

BRAKMAN, S., J. H. GARRETSEN, AND C. VAN MARREWIJK (2009): "Economic Geography within and between European Nations: The Role of Market Potential and Density across Space and Time," *Journal of Regional Science*, 49(4), 777–800.

BRAUNERHJELM, P. (2008): "Specialization of Regions and Universities: The New Versus the Old," *Industry and Innovation*, 15(3), 253–275.

CARLTON, D. W. (1983): "The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables," *Review of Economics and Statistics*, 65(3), 440–449.

CHRISTOPOPOULOS, D. K., AND E. G. TSIONAS (2007): "Are Regional Incomes in the USA Converging? A Non-linear Perspective," *Regional Studies*, 41(4), 525–530.

CRONE, T. M. (2005): "An Alternative Definition of Economic Regions in the United States Based on Similarities in the State Business Cycle," *Review of Economics and Statistics*, 87(4), 617–626.

CRONE, T. M., AND A. CLAYTON-MATTHEWS (2005): "Consistent Indices for the 50 States," *Review of Economics and Statistics*, 87(4), 593–603.

ELLICKSON, P. B., S. HOUGHTON, AND C. TIMMINS (2010): "Estimating Network Economies in Retail Chains: A Revealed Preference Approach," Working Paper No. 15832, National Bureau of Economic Research, Cambridge, MA.

ELLISON, G., AND E. L. GLAESER (1997): "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *The Journal of Political Economy*, 105(5), 889–927.

——— (1999): "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration," *American Economic Review*, 89(2), 311–316.

ELLISON, G., E. L. GLAESER, AND W. KERR (2010): "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 100(1), 1–42.

FALCK, O., M. FRITSCH, AND S. HEBLICH (2009): "Bohemians, Human Capital, and Regional Economic Growth," CESifo Working Paper No. 2715, Münchener Gesellschaft zur Förderung der Wirtschaftswissenschaften, Munich.

FLORIDA, R. (2002a): "Bohemia and Economic Geography," *Journal of Economic Geography*, 2, 55–71.

——— (2002b): *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*. Basic Books, New York, NY.

——— (2005): "The World Is Spiky: Globalization Has Changed the Economic Playing Field, But Hasn't Leveled It," *Atlantic Monthly*, 147(9), 48–53.

FLORIDA, R., G. GATES, B. KNUDSEN, AND K. STOLARICK (2006): "The University and the Creative Economy," Mimeograph, CreativeClass.com.

FLORIDA, R., T. GULDEN, AND C. MELLANDER (2008): "The Rise of the Mega-Region," *Cambridge Journal of Regions, Economics and Society*, 1(3), 459–476.

FLORIDA, R., C. MELLANDER, AND K. STOLARICK (2008): "Inside the Black Box of Regional Development – Human Capital, the Creative Class and Tolerance," *Journal of Economic Geography*, 8(3), 615–649.

FOUSEKIS, P. (2007): "Convergence of Relative State-level Per Capita Incomes in the United States Revisited," *Journal of Regional Analysis and Policy*, 37(2), 80–89.

FRALEY, C., AND A. E. RAFTERY (2002): "Model-Based Clustering, Discriminant Analysis and Density Estimation," *Journal of the American Statistical Association*, 97(458), 611–631.

——— (2006): "MCLUST Version 3 for R: Normal Mixture Modeling and Model-Based Clustering," Technical Report 504, University of Washington, Department of Statistics.

FRIEDMAN, T. L. (2005): *The World Is Flat: A Brief History of the Twenty-first Century*. Farrar, Straus and Giroux, New York, NY, 1 edn.

GLAESER, E. L. (1999): "Learning in Cities," *Journal of Urban Economics*, 46(2), 254–277.

GRILICHES, Z. (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," *Bell Journal of Economics*, 10(1), 92–116.

GUIMARÃES, P., O. FIGUEIREDO, AND D. WOODWARD (2003): "A Tractable Approach to the Firm Location Decision Problem," *Review of Economic Studies*, 85(1), 201–204.

GUIMARÃES, P., AND R. C. LINDROOTH (2007): "Controlling for Overdispersion in Grouped Conditional Logit Models: A Computationally Simple Application of Dirichlet-multinomial Regression," *Econometrics Journal*, 10(4), 439–452.

HAMERMESH, D. S. (2007): "Viewpoint: Replication in Economics," *Canadian Journal of Economics*, 40(3), 715–733.

HECKER, D. E. (2005): "High-technology employment: a NAICS-based update," *Monthly Labor Review*, 128(7), 57–70.

HELLERSTEIN, D., AND R. MENDELSOHN (1993): "A Theoretical Foundation for Count Data Models," *American Journal of Agricultural Economics*, 75(3), 604–611.

HIRSCH-KREINSEN, H. (2008): ""Low-Tech" Innovations," *Industry and Innovation*, 15(1), 19–43.

HOYMAN, M., AND C. FARICY (2009): "It Takes a Village: A Test of the Creative Class, Social Capital, and Human Capital Theories," *Urban Affairs Review*, 44(3), 311–333.

JAFFE, A. B., M. TRAJTENBERG, AND R. HENDERSON (1993): "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108(3), 577–598.

KOO, J. (2007): "Determinants of Localized Technology Spillovers: Role of Regional and Industrial Attributes," *Regional Studies*, 41(7), 995–1011.

LEE, S. Y., R. FLORIDA, AND Z. J. ACS (2004): "Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation," *Regional Studies*, 38(8), 879–891.

LUCAS, ROBERT E., J. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22(1), 3–42.

MANKIW, N. G., D. ROMER, AND D. N. WEIL (1992): "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107(2), 407–437.

MAYER, H. (2009): "Bootstrapping High-Tech: Evidence from Three Emerging High Technology Metropolitan Areas," Metropolitan Economy Initiative No. 10, Brookings Institution, Washington, DC.

MCCANN, P. (2008): "Globalization and Economic Geography: The World is Curved, not Flat," *Cambridge Journal of Regions, Economics and Society*, 1(3), 351–370.

MCFADDEN, D. (1974): *Frontiers in Econometrics*chap. Conditional Logit Analysis of Qualitative Choice Behaviour, pp. 105–142. Academic Press, New York.

MCGRANAHAN, D. A., AND T. R. WOJAN (2007): "Recasting the Creative Class to Examine Growth Processes in Rural and Urban Counties," *Regional Studies*, 41(2), 197–214.

MCGRANAHAN, D. A., T. R. WOJAN, AND D. M. LAMBERT (2010): "The Rural Growth Trifecta: Outdoor Amenities, Creative Class and Entrepreneurial Context," *Journal of Economic Geography*, forthcoming.

O'HUALLACHÁIN, B. (2008): "Regional Growth Transition Clubs in the United States," *Papers in Regional Science*, 87(1), 33–54.

PANEK, S. D., F. T. BAUMGARDNER, AND M. J. MCCORMICK (2007): "Introducing New Measures of the Metropolitan Economy: Prototype GDP-by-Metropolitan-Area Estimates for 2001?-2005," *BEA Survey of Current Business*, 87(11), 79–113.

PORTER, M. E. (2000): "Location, Competition, and Economic Development: Local Clusters in a Global Economy," *Economic Development Quarterly*, 14(1), 15–34.

SALA-I MARTIN, X. (2006): "The World Distribution of Income: Falling Poverty ... and Convergence, Period," *Quarterly Journal of Economics*, 121(2), 351–397.

SAXENIAN, A. (1994): *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.

SCHERER, F. M. (1982): "Inter-Industry Technology Flows and Productivity Growth," *Review of Economics and Statistics*, 64(4), 627–634.

SHAPIRO, J. M. (2006): "Smart Cities: Quality of Life, Productivity and the Growth Effects of Human Capital," *Review of Economics and Statistics*, 88(2), 324–335.

SHEARMUR, R., AND D. DOLOREUX (2009): "Place, Space and Distance: Towards a Geography of Knowledge-Intensive Business Services Innovation," *Industry and Innovation*, 16(1), 79–102.

STORPER, M., AND A. J. SCOTT (2009): "Rethinking Human Capital, Creativity and Urban Growth," *Journal of Economic Geography*, 9(2), 147–167.

TIMMINS, C. (2005): "Estimable Equilibrium Models of Locational Sorting and Their Role in Development Economics," *Journal of Economic Geography*, 5(1), 83–100.

WOJAN, T. R., D. M. LAMBERT, AND D. A. MCGRANAHAN (2007): "Emoting with Their Feet: Bohemian Attraction to Creative Milieu," *Journal of Economic Geography*, 7(5), 711–736.

WOODWARD, D., O. FIGUEIREDO, AND P. GUIMARÃES (2006): "Beyond the Silicon Valley: University R&D and High-Technology Location," *Journal of Urban Economics*, 60(2), 15–32.

YOUNG, A. T., M. J. HIGGINS, AND D. LEVY (2008): "Sigma Convergence versus Beta Convergence: Evidence from U.S. County-Level Data," *Journal of Money, Credit and Banking*, 40(5), 1083–1093.

Chapter 3

Accounting for amenities: Evidence on expenditures for non-market goods and services that influence the quality of life[☆]

with Nicolai V. Kuminoff and Jaren C. Pope

3.1 Abstract

Tracking the value of non-market goods and services that affect the quality of life is important to households, businesses and policymakers alike. However, while the National Income and Product Accounts (NIPA) fail to account for amenity values, there is widespread evidence that spatial variation in rents and wages reflects the values that households assign to environmental services, public goods, and urban activities. Accurate estimates for personal consumption expenditures on quality of life would complement the corresponding NIPA measures for private goods. Popular rankings are unsuitable to track amenity values because they are typically not grounded in economic theory. Yet many theoretically consistent quality-of-life rankings that derive hedonic wage and rent differentials lack the intuitive appeal of the

[☆] Jointly authored with Nicolai V. Kuminoff (Dept. of Economics, Arizona State University) and Jaren C. Pope (Dept. of Economics, Brigham Young University). I am the lead author and responsible for the main text and empirical work.

former. The problem is that the best estimates to date suffer from an index number analog to the classic omitted variable problem. This helps to explain why past estimates often seem counterintuitive. We revisit these issues and provide the first large-scale update of the seminal study of county-level quality-of-life rankings by Blomquist, Berger, and Hoehn (1988, BBH henceforth). We significantly expand the set of amenities used in the previous literature by introducing broad categories of geographical, environmental, neighbourhood, infrastructure and urbanisation amenities. This paper increases the geographic coverage beyond urban areas by producing rankings for the entire United States. We also introduce a number of methodological improvements to the literature. A more accurate measure of imputed rents is presented which can account for the large spatial variation in the user cost of housing.
Keywords: Quality of life, national income and product accounts, amenities, public goods, hedonic
JEL Codes: E01, H41, Q51, Q56, R12, R13

3.2 Introduction

It has been argued that when a household chooses a place to live, it makes "the most important decision of life" (Florida, 2008). Indeed, on average close to 30 million such decisions were made every year in the United States over the last decade, with one third of re-locations taking place across state lines. Given such high levels of mobility, it is perhaps not surprising that the ranking of U.S. localities according to specific quality-of-life criteria has had a long tradition: From Sperling's *Cities Ranked & Rated* to the *PlacesRates Almanac*, Money's *Best Places to Live*, Kiplinger's *Best Cities*, Forbes *Best Places for Business and Careers*, relative differences in the quality of life do not only capture the attention of households, but they are also closely monitored by firms and local policy makers.[1] But because it is often unclear how various quality-of-life factors are weighted or because the weighting scheme merely reflects the index provider's own preferences, popular rankings remain largely without a firm theoretical anchor which renders them unsuitable for policy purposes. By contrast, the weights of rankings that are consistent with economic theory would only reflect the market and non-market preferences that are revealed by households' labour and housing market choices. Beyond simply producing rankings, valuing non-market amenities has the increasingly important use of deriving willingness-to-pay measures for goods that are not included in national income accounts or standard cost-of-living indices.

Furthermore, traditional measures of economic welfare do not fully reflect the impact of urbanisation and agglomeration. Yet without these residential and occupational transformations much of the

[1] Recently, a number of global quality-of-life rankings have been publicised as compensation management tools for multinational corporations or as inputs for local economic development efforts that target global jobs. The Economist's Intelligence Unit's rankings for countries (EIU, 2005) and Mercer's rankings for global cities (Mercer, 2009) are the most prominent global such league tables.

technological progress since the Industrial Revolution could simply not have been converted into economic growth. Although calls for more comprehensive welfare measures go back to Nordhaus and Tobin (1972), the growing importance of non-market assets such as the environment together with the availability of unprecedented high-resolution data has renewed broad interest in "green accounting" (see e.g. Nordhaus, 2000; Banzhaf, 2005). In order to understand what portion of the higher earnings of urban residents may simply be compensation for the negative externalities of urbanisation, policy makers need more reliable measures of how households' locational choices are affected by the non-price interaction of non-market goods.

Urban economists generally think about quality of life in terms of the relative importance of different factors to household well-being, usually expressed as utility. The key insight of this literature rests on the observation that location-specific differences in wages and (land) rents should compensate for the differences in non-market characteristics, such as natural or cultural amenities that increase the attractiveness of a given locality. Thus, although geographic disparities in the quality of life themselves are unobservable, they can be measured by prevailing wage differentials with the local cost-of-living netted out. Since the work of Rosen (1979), Haurin (1980) and Roback (1982, 1988), a growing body of literature has tried to produce theoretically consistent quality-of-life rankings for urban areas by deriving wage and rent differentials via hedonic methods, calculating the implicit prices of location-specific amenities which are then used as utility valuation weights.[2]

While this early literature focuses on interurban differences and treats cities as spatially homogenous entities, extensions by Hoehn, Berger, and Blomquist (1987, HBB henceforth) incorporate urban structure and intraurban location into the quality-of-life framework and allow for amenity variation both within and across urban areas. In addition to spatial structure, the augmented Roback-HBB framework also incorporates city-size related agglomeration dynamics via productivity effects, establishing key linkages between the firms in a given location. In this setting, the relationship between quality of life and urban size remains an empirical matter, ultimately determined by the precise nature of these agglomeration effects and the relative impact of (dis)amenities on production and consumption. If a household amenity also reduces unit production costs of firms (e.g. public infrastructure), the comparative statics of the Roback-HBB model imply that the quality of life unambiguously increases with city size, irrespective of whether the agglomeration effects have a positive or negative impact on productivity. A string

[2]See, for example, Gyourko, Kahn, and Tracy (1999), Blomquist (2006) and Lambiri, Biagi, and Royuela (2007) for comprehensive surveys of the literature.

of prominent papers – including Blomquist, Berger, and Hoehn (1988), Gyourko and Tracy (1991) and, most recently, Chen and Rosenthal (2008) – produce quality-of-life rankings that are negatively related to city size.[3] At a minimum, this research suggests that the bulk of amenities that contribute to quality of life is at best unproductive from the firm's perspective, possibly even pointing to the presence of productivity-reducing congestion effects. Put differently, households and firms do not appear to have the same preferences for city size.[4]

However, this evidence seems difficult to reconcile with the popular suggestion that the quality of life is an increasingly important driver of urban development patterns, a claim that rests on the hypothesis that the demand for consumption amenities has grown as incomes and education levels have risen nationwide (Glaeser, Kolko, and Saiz, 2001). Indeed, while firm productivity advantages continue to dominate amenities in most metropolitan areas, the relative importance of amenities appears to have increased between 1990 and 2000 (Deitz and Abel, 2008).[5] It is thus perhaps not surprising that existing hedonic quality-of-life rankings have been faulted for their lack of intuitive appeal and their tendency to produce apparent misrankings (Rappaport, 2008). In the best rankings available to date, some highly ranked counties have high rates of poverty and crime (e.g. Bibb, GA; Kern, CA). Some low ranked counties are famous for their scenic beauty, good schools, low crime, and high incomes (e.g. Marin, CA; Westchester, NY).

The problem is that these estimates suffer from an index number analog to the classic omitted variable problem. This helps to explain why past estimates often seem counterintuitive, a criticism that is compounded by the fact that past rankings bear no resemblance to the quality-of-life rankings produced in the popular press.[6] Furthermore, focusing on inter-urban quality-of-life differences alone, the current literature does not capture the vast array of intra-urban diversity among US cities. Large differences in urban form notwithstanding, the benchmark study of BBH remains the only comprehensive set of urban rankings to date with sufficient spatial resolution to document evidence of both inter- and intraurban variation in the quality of life for metropolitan areas in the United States.

---

[3] Albouy (2008, 2009b) is a notable exception in this regard. His methodological improvements create metro-level rankings that increase with city size by incorporating a number of adjustments into the basic Roback-HBB framework, specifically by accounting for federal taxes, non-labour income and non-housing cost. We discuss how this work relates to ours below.

[4] This is indeed one of the key insights in Gabriel and Rosenthal (2004) and Chen and Rosenthal (2008).

[5] Local productivity does not have to be the main driver of urbanisation. Rappaport (2009) shows that the growing importance of quality-of-life differentials can eventually cause local amenities to become the sole determinant of relative local density.

[6] The correlations with the rankings of the 1981 PlacesRated Almanac are -0.1795 for Roback (1982), -0.0443 for Blomquist, Berger, and Hoehn (1988), -0.0508 for Gyourko and Tracy (1991) and – compared to the 2007 PlacesRated Almanac – the corresponding correlation for Chen and Rosenthal (2008) is -0.0257.

The purpose of this paper is to address the limitations in the quality-of-life literature in three major ways: (i) We revise the current methodology for estimating quality of life expenditures; (ii) we present new estimates for U.S. counties; and (iii) we relate our estimates to NIPA measures for expenditures on private goods. We make several improvements to the methodology developed by Blomquist, Berger, and Hoehn (1988, BBH henceforth). First, we assemble the most comprehensive data on amenities to date. In addition to the sixteen variables collected by BBH (e.g. sunshine, precipitation, visibility, violent crime) we track an additional 60 amenities, ensuring a more nuanced differentiation in terms of how locations vary with regard to the underlying determinants of the quality of life. The expanded set of amenities thus minimises the omitted variable bias of the aggregate index and permits a more accurate attribution of what portion of the spatial variation in implicit quality-of-life expenditure on are due to local differences in climate and geography, environmental externalities, local public goods, infrastructure and urbanisation amenities. Second, we control for spatial variation in the user cost of housing, purchasing power, tax burden, and non-wage income (Albouy, 2008).

Given the pivotal role of the cost of housing in determining location-specific quality-of-life estimates, we show that a more accurate measure for housing expenditure has a more prominent impact on quality-of-life estimates than other refinements. Third, we improve on the BBH and Kahn (1995) estimators using a two-step approach based on Evans and Smith (2005), avoiding some of the empirical pitfalls of earlier studies. Fourth, we extend the scope of the analysis from the 253 counties in BBH to all metropolitan counties and then to every county in the contiguous United States. Extending our analysis using the 1999 OMB definition of metropolitan areas (1,085 counties) increases the urban population coverage from 47% to 80% and expands the urban land coverage from 9% to 30%. Finally, we use the largest possible source of micro dataŮthe 5% PUMS sample from the 2000 Census. Thus, our regressions use data from approximately 10 million consumers.

Our results indicate major differences between the updated rankings and those presented by BBH. Under a variety of alternative specifications, the rank correlations between our rankings and the BBH rankings range from 0.09 to 0.14. Furthermore, the new rankings are stable across our geographical samples and do not change as the geographical coverage is increased to span the full rural-urban spectrum of counties. Irrespective of whether we use the restricted set of original amenities or the expanded set of new amenities, the new rankings tend to increase with city size which is consistent with the tenor of popular rankings that big cities are not necessarily bad places to live. More importantly from a theoretical perspective, our results point towards the increasing importance of consumption amenities as the willingness to pay for quality of life increases with population density. These findings are substantiated further by the full implicit price estimates across different amenity categories. Moreover, the new rankings are also consistent with the notion that quality of life has become an increasingly important consumption item as measured by its implied budget share. Our rankings indicate a willingness to pay between 8% to 24% of household income in order to live in the ten most desirable locations in the nation, compared to only 3% to 9% reported in BBH.

The remainder of the paper proceeds as follows. Section 3.3 provides a brief summary of the Roback-HBB framework in which we ground the empirical analysis for our quality-of-life estimates. The econometric methodology based on a two-stage hedonic quality-of-life estimator is presented in section 3.4, followed by section 3.5 which describes our data set on wages, housing expenditure and discusses the selection of the different categories of amenities. Our hedonic estimates are presented in section 3.6. Section 3.7 then discusses the new rankings in some detail, whereas section 3.8 relates our estimates to NIPA measures for expenditures on private goods and lays the foundation for developing a consistent macroeconomic index of amenity value in the United States. Section 3.9 concludes the paper.

3.3 Rosen-Roback model

In the standard Rosen-Roback model implicit amenity prices follow from wage and rent (housing expenditure) differentials in a dual-market sorting equilibrium. The national (closed) economy is characterised by a finite number of spatially-bounded localities (e.g. counties) that differ with regard to their specific combination of wages w and the cost of housing or rents r. Places can also vary in terms of the prevailing level of quality of life q, where q_j is some index of the area j-specific bundle of K amenities, $a_{k\in K,j}$, such that $q_j = \phi(\boldsymbol{A}_j)$ and $\boldsymbol{A}_j = [a_{1j}, \ldots, a_{kj}]$ for $j = 1, 2, \ldots, J$. The national economy is divided into rural and urban areas where an arbitrary metropolitan region is composed of one or several localities. Across these sites, fully mobile households maximise their well-being subject to a budget constraint and footloose firms minimise their cost by making their respective location choices.

In site j, homogenous households enjoy the local quality of life q_j and consume a composite commodity that consists of a nationally-traded good x with price p^x which will be taken as numéraire and local, non-tradable housing services h_j which cost r^h_j. Households are assumed to live and work in the same location where they receive income from supplying one unit of labour at the local wage rate w_j. In addition to wage income, total household income $m_j = w_j + I$ also consists of a locally invariant component of non-wage income I, which gives rise to the following choice problem:

$$\max_{j,x,h} \; U(x,h;q_j) \qquad \text{subject to} \quad w_j = x + r_j^h h, \tag{A-1}$$

$$\text{where} \quad q_j = \theta_1 a_{1j} + \theta_2 a_{2j} + \ldots + \theta_k a_{kj}. \tag{A-2}$$

The individual indirect utility function can now be written as $V_j = v(w_j, r_j^h; q_j)$. Similarly, firms minimise costs $C_j = c(w_j, r_j^b; \tilde{q}_j)$, where $\tilde{q}_j = \varphi(\boldsymbol{A}_j)$ and $\varphi(\cdot) \neq \phi(\cdot)$.[7] For simplicity, firms and households are assumed to incur the same amount of quality-adjusted rental expenditure per unit of space, i.e. $r_j^h = r_j^b = r_j$. The set of household and firm optimisation decisions that sustains equilibrium in site j can thus be expressed as:

$$\overline{V} \quad = \quad v(w_j, r_j; q_j), \quad \forall \quad j = 1, \ldots, J \quad \text{and} \tag{A-3}$$

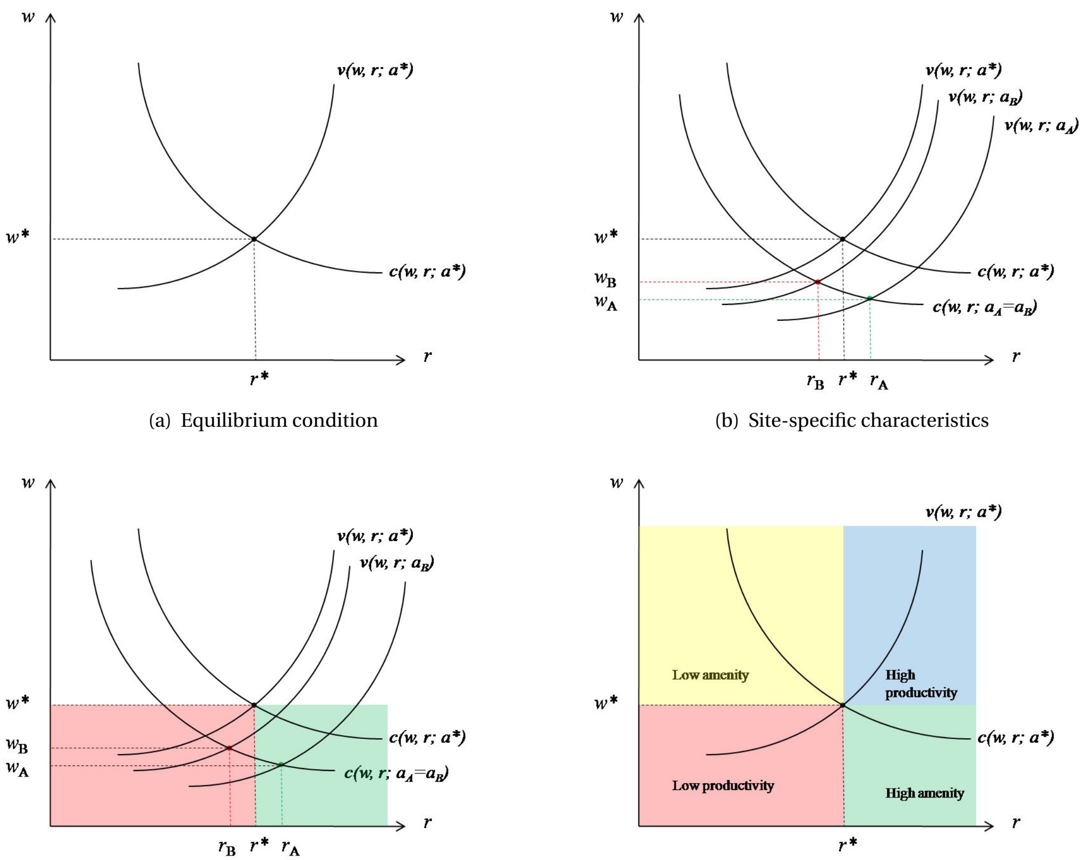
$$\overline{C} \quad = \quad c(w_j, r_j; \tilde{q}_j). \tag{A-4}$$

Locational equilibrium implies identical levels of well-being and costs across locations as households cannot improve their utility and firms cannot reduce their total costs by relocating. In wage-rent space, equilibrium is therefore characterised by the average isoutility and isocost curves that determine equilibrium wages w^* and rents r^* as illustrated in figure 3.1(a). The comparative statics of location-specific differences in amenities is illustrated in figure 3.1(b), where location A has above-average amenities and below average productivity and location B is also endowed with above-above average amenities ($q^* < q_B \leq q_A$), but the productivity effect dominates the amenity effect in location B. Thus, the net effect of location-specific amenity differentials on rents depends on the relative shift of the isoutility and isocost curves. In turn, the make-up of these (observable) differentials permits inference on the nature of a location's (unobservable) productivity and amenity characteristics in figure 3.1(d)

In this setting, equilibrium wage and rent differentials are used to compute the implicit prices of amenities since utility opportunities within each location are equal in equilibrium. Taking the total dif-

[7] In the context of firms, $\tilde{q}$ can be thought of in terms of the "quality of business" environment, i.e. the additional cost that a firm is willing to incur in order to operate in a location with a specific set of attributes. This idea was first introduced by Gabriel and Rosenthal (2004).

Figure 3.1: Mechanics of the Rosen-Roback general equilibrium model



(a) Equilibrium condition

(b) Site-specific characteristics

(c) Amenity and productivity effects

(d) Different dominant effects

Notes: Figure 3.1(a) illustrates the wage-rent equilibrium. Figure 3.1(b) shows the impact of a differential amenity effect for two locations, *A* and *B*, that experience the same productivity effect. Both locations *A* and *B* have above average amenities (lower wages), but as figure 3.1(c) illustrates in *A* the the amenity effect dominates (higher rent) and in *B* the productivity effect dominates (lower rent). Different amenity endowments relative to the average location in combination with the relative strength of productivity and amenity effects support different combinations of inter-urban wage-rent differentials. In turn, the make-up of these (observable) differentials permits inference on the nature of a location's (unobservable) productivity and amenity characteristics in figure 3.1(d).

ferential of the representative household's indirect utility function V_j in equation (A-3), this implies that

$$dV = 0 = \frac{\partial v}{\partial w} dw_j + \frac{\partial v}{\partial r} dr_j + \frac{\partial v}{\partial q}\frac{\partial q}{\partial a} da_{kj}. \tag{A-5}$$

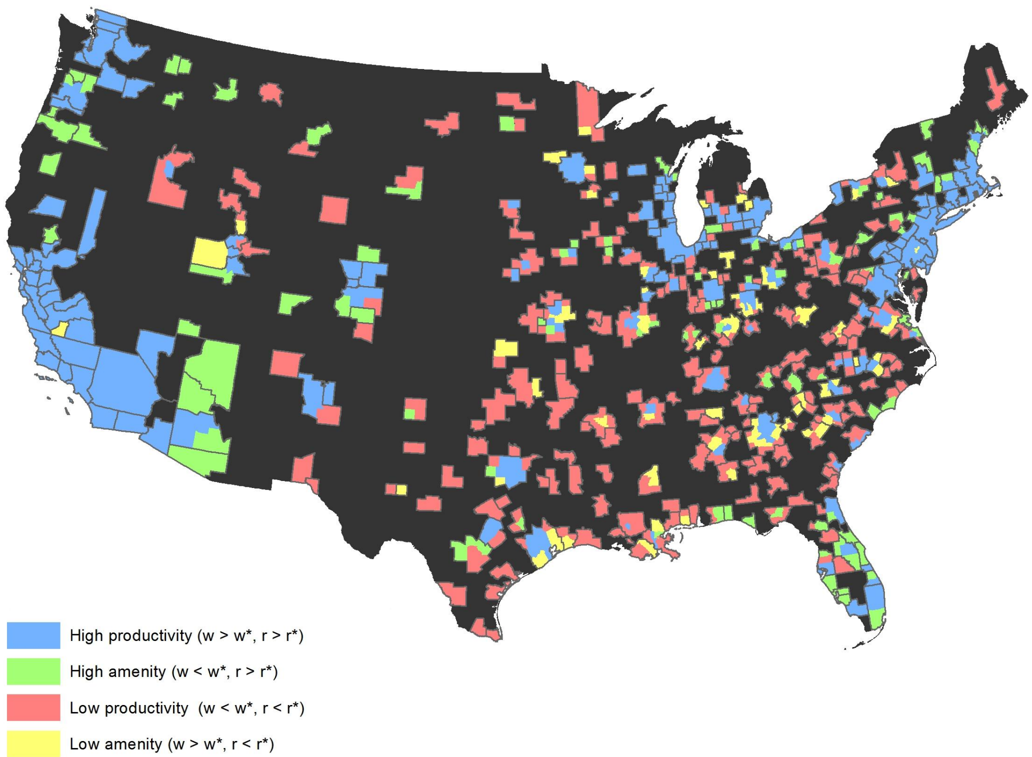
We can now analyse the household's willingness to pay for an additional unit of a single amenity a_k (e.g. scenic beauty or live performance venues), holding constant all other local attributes. Let the implicit price for this amenity a_k be defined as $p^{a_k} = \frac{\partial v}{\partial q}\frac{\partial q}{\partial a} / \frac{\partial v}{\partial w}$. By rearranging equation (A-5) using Roy's identity, we have Roback's (1982) result that the implicit amenity price is equal to the housing expenditure differential minus the wage differential:

$$p_j^{a_k} \quad = \quad h_j(dr_j/da_{kj}) - dw_j/da_{kj}, \tag{A-6}$$

where h^j is the amount of housing purchased by a household in location j. Empirically, $p_j^{a_k}$ is measured via the hedonic gradients, dr_j/da_{kj} and dw_j/da_{kj}, that are retrieved from housing and wage regressions respectively. The K individual equilibrium amenity prices are then aggregated into site-specific quality-of-life indices p_j^q. The econometric methodology of deriving the quality-of-life indices from hedonic estimates of implicit amenity prices is discussed next.

Figure 3.2 illustrates which effect dominates across the metropolitan areas in the contiguous United States. While the sizeable productivity advantages of the largest metro areas are clearly visible, the granularity of our county-level data also shows that the traditional core-periphery productivity gradients of the monocentric city models appear to hold true for a number of urban areas. In particular, for cities such as Atlanta, Cincinnati, Dallas-Fort Worth, Pittsburg, Richmond, or St. Louis, the high-productivity central business district and the concentric arrangement of low amenity and low productivity areas around it are clearly visible. Similarly, high amenity areas in these cities either tend to be in the mature suburbs that are immediately adjacent to the CBD or in traditional, outlying commuter communities. In what follows, our quality-of-life metrics will capture how much of this variation can be attributed to local disparities in the bundle of amenity endowment.

Figure 3.2: Dominant amenity and productivity effects for metropolitan counties



Notes: Among the 359 metropolitan statistical areas in the conterminous United States, 51% of counties have above average wages and rents and are labelled "high productivity" (i.e. $w > w^*$ and $r > r^*$, where w^* and r^* are the household wages and housing expenditure for the metropolitan average). 14% of these metropolitan counties are "high amenity" ($w < w^*$, $r > r^*$), whereas 22% are "low productivity" ($w < w^*$, $r < r^*$), and 13% are "low amenity" ($w > w^*$, $r < r^*$).

3.4 Econometric methodology

Our strategy for the overall empirical analysis broadly follows the logic of Evans and Smith (2005) and Combes, Duranton, and Gobillon (2008). We derive an empirical measures for the price of quality of life p_j^q that is based on the equilibrium condition in equation (A-6) in two stages. With sufficient mobility across labour and housing markets, these market-determined implicit prices provide valid estimates of a household's marginal willingness to pay for specific amenities. Conceptually, the first stage identifies the hedonic vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ from individual housing expenditure and wage functions, conditional on a set of individual and structural characteristics $\boldsymbol{X}_{ij}$ and the local amenities $\boldsymbol{A}_j$:

$$\left.\begin{aligned} \text{Housing expenditure function:}\quad & r_{ij} = f(\boldsymbol{X}_{ij}^r, \boldsymbol{A}_j; \boldsymbol{\beta}) \\ \text{Wage function:}\quad & w_{ij} = g(\boldsymbol{X}_{ij}^w, \boldsymbol{A}_j; \boldsymbol{\gamma}) \end{aligned}\right\}\text{where}\quad \boldsymbol{A}_j = [a_{1j}, \ldots, a_{kj}]. \tag{A-7}$$

In the second stage, location-specific prices of quality of life p_j^q are then derived as a weighted average of all K amenities prevailing in location j, using implicit equilibrium prices as appropriate valuation weights given their consistency with household utility maximisation:

$$\begin{aligned} \text{QOLI}_j \equiv p_j^q \quad &= \quad \sum_{k=1}^{K} a_{kj} p_j^{a_k} \quad \text{with} \quad j = 1, 2, \ldots, J \\ &= \quad \sum_{k=1}^{K} a_{kj} \left[\frac{\partial r}{\partial a_k}(\boldsymbol{X}_{ij}^r, \boldsymbol{A}_j; \hat{\boldsymbol{\beta}}) - \frac{\partial w}{\partial a_k}(\boldsymbol{X}_{ij}^w, \boldsymbol{A}_j; \hat{\boldsymbol{\gamma}}) \right]. \end{aligned} \tag{A-8}$$

Empirically, we first estimate hedonic regressions with area-specific fixed effects for both the wage and housing expenditure function.[8] The estimated location-specific fixed effects are then used along with the estimated covariance matrix in a second-stage, feasible generalized least squares (FGLS) regression. This model evaluates the marginal impact of individual amenities in each locality on the estimated fixed effects, thus providing the hedonic gradients required for the estimation of full implicit amenity prices $p_j^{a_k}$.

[8] Our use of location dummies to measure local wage and rent differences is also related to the approach taken by Gabriel and Rosenthal (2004), Shapiro (2006) and Chen and Rosenthal (2008). In contrast to their work, we do not directly compute quality-of-life measures from these fixed effects which capture both observable and unobservable local characteristics, but instead we investigate how much of the variation of these dummies can be explained by the variation in an explicit set of the location-specific amenities.

3.4.1 First-stage hedonic estimation

In order to account for local differences in amenities and other location-specific effects, the first-stage hedonic housing and wage regressions include spatial fixed effects, λ_j^r and λ_j^w respectively, such that:

$$\begin{aligned} r_{ij} &= \boldsymbol{X}_{ij}^r \boldsymbol{\beta} + \lambda_j^r + \varepsilon_{ij} && \text{(A-9)} \\ w_{ij} &= \boldsymbol{X}_{ij}^w \boldsymbol{\gamma} + \lambda_j^w + \nu_{ij}, && \text{(A-10)} \end{aligned}$$

where r_{ij} are log individual monthly housing expenditures, w_{ij} are log hourly wages, and $\boldsymbol{X}_{ij}^{(r,w)}$ is a vector of structural housing characteristics and individual attributes respectively. The location-specific intercept $\lambda_j^{(r,w)}$ measures observable and unobservable factors that may contribute to the local variation in wages and housing expenditures, but that do not vary among residents (and their homes) within each location. Estimates of these fixed effects $\hat{\lambda}_j^{(r,w)}$ are used as the dependent variable in the second stage in order to retrieve the marginal effects for each amenity.

3.4.2 Second-stage estimation of quality-of-life index

The spatial fixed-effects are thought to be a function of location-specific amenities such that $\lambda_j = \phi(\boldsymbol{A}_j)$, which then yields a wage and a housing second-stage estimating equation of the following form:

$$\left[\hat{\lambda}_j^r, \hat{\lambda}_j^w\right] = \hat{\lambda}_j = \alpha + \boldsymbol{A}_j \boldsymbol{\delta} + \xi_j \tag{A-11}$$

where $\boldsymbol{A}_j$ represents the vector of local amenities.[9] Suppressing the subscripts and summarizing the model in matrix form, estimates for the amenity coefficients, $\boldsymbol{\delta}$, can be derived using an FGLS estimator as in equation (A-12):

$$\hat{\boldsymbol{\delta}} = (\boldsymbol{A}'\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{A})^{-1}\boldsymbol{A}'\hat{\boldsymbol{\Omega}}^{-1}\hat{\lambda} \tag{A-12}$$

[9]Since $\hat{\lambda}_j$ is a dummy variable coefficient estimated from semilogarithmic equations in the first stage, we use the empirical implementation of the Halvorsen-Palmquist adjustment proposed by Kennedy (1981) to create the dependent variables for the second stage.

where $\hat{\boldsymbol{\Omega}} \equiv \text{VCOV}(\hat{\lambda})$, the estimated covariance matrix associated with the spatial fixed effects, is obtained from the first-stage model. Following equation (A-6), the vector of implicit amenity prices is then calculated as $\hat{\boldsymbol{p}}^a = h\hat{\boldsymbol{\delta}}^r - \hat{\boldsymbol{\delta}}^w$ and the quality-of-life estimates are given by $\hat{\boldsymbol{p}}^q = \boldsymbol{A}\hat{\boldsymbol{p}}^a$.

Equations (A-9) to (A-12) represent the full econometric specification of our model. An alternative specification would have been to perform a single-stage estimation and use all the explanatory variables at once. However, using a single-stage framework might be problematic for a variety of reasons. First, the one-stage estimation does not permit the computation of the variance of location-specific local shocks, ξ_j in equation A-11. At the same time, we cannot distinguish aggregate local shocks from purely idiosyncratic shocks in the microdata at the household level, which is important due to the presence of unobserved characteristics. Furthermore, in a single-stage estimation, we would try to measure the effect of aggregate variables – such as local amenities – on microdata by merging the aggregate data with micro observations. Moulton (1990) cautions that this single-stage methodology creates large biases in the standard errors for the estimated coefficients of aggregate explanatory variables, because of the unrealistic assumption of independent disturbances for the aggregate and microdata, respectively. The two-stage estimation method employed here avoids these problems. In order to check the robustness of our results, we ran a single-stage estimation and found qualitatively similar results for the estimated coefficients.

3.4.3 Omitted variable bias in QOLI estimation

Estimates for expenditures on quality of life potentially suffer from an index analog to the classic omitted variable problem. In order to understand why past estimates often seem counterintuitive, it is useful to consider hedonic quality of life estimation as an atypical omitted variable problem with the goal to recover the composite index QOLI_j, rather than obtaining precise point estimates of marginal prices of individual amenities $\hat{\delta}$. Omitted variable bias affects the aggregate index estimates in a slightly different way than it distorts individual amenity price estimates, since

$$\text{QOLI}_j = \sum_{k=1}^{K} a_{kj}\left(\hat{\delta}_k^r - \hat{\delta}_k^w\right), \quad \text{where} \quad \hat{\lambda}_j = \alpha + \boldsymbol{A}_j\boldsymbol{\delta} + \xi_j. \tag{A-13}$$

These single hedonic prices are only fully immune to omitted variable bias in the unlikely event that no location- or household-specific unobservables are correlated with the observable amenities. As a consequence, any correlation between observed and unobserved amenities will bias the hedonic estimates for the marginal amenity prices.[10] In the case of the quality-of-life index, however, high correlation between observed and unobserved amenities attenuates the omitted variable bias of the aggregate index. Less than perfect correlation between observed and unobserved amenities will lead to biased estimates for the aggregate quality-of-life indices. This implies that, while the marginal benefit from including additional data on amenities that are highly correlated with existing amenities is comparatively small, including weakly correlated amenities potentially has dramatic effects on the overall quality-of-life estimate. The next section illustrates this formally and provides an empirical example that quantifies the magnitude of the omitted variable bias index effect.

A simple example

Let $K = 1$ in equation (A-11) and assume that, in addition to the only observed amenity a_j with implicit price δ, location j is also characterised by an unobserved amenity z_j with implicit price κ, such that

$$\hat{\lambda}_j = \alpha + a_j\delta + \xi_j = \alpha + a_j\delta + (z_j\kappa + \zeta_j). \tag{A-14}$$

We are interested in obtaining an unbiased estimate of the true quality-of-life index $\text{QOLI}_j = a_j\delta + z_j\kappa$. However, given that amenity z_j is unobservable (to the econometrician), we instead identify

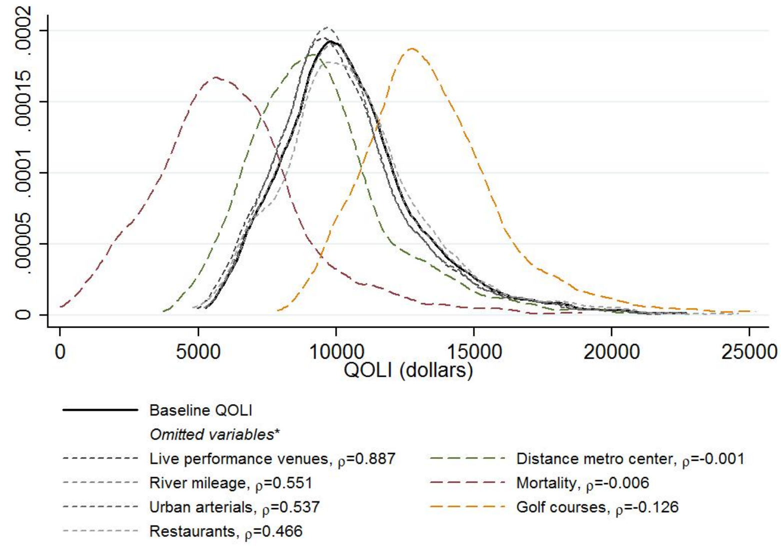
$$\widetilde{\text{QOLI}}_j = a_j\tilde{\delta} = a_j\delta + (z_j - \eta_j)\kappa, \tag{A-15}$$

where $z_j = a_j\pi + \eta_j$ and $\text{E}(\eta_j|a_j) = 0$. Since $\pi = \rho_{az}\frac{\sigma_z}{\sigma_a}$, where ρ_{az} is the correlation between a_j and z_j, and the sigmas are the corresponding standard deviations, equation (A-15) can be restated as $\widetilde{\text{QOLI}}_j = a_j\left(\delta + \left[\rho_{az}\frac{\sigma_z}{\sigma_a}\right]\kappa\right)$. We can now express our earlier observation regarding the index analog of the omitted variable bias associated with unobserved amenities more formally. First, as $\pi \to 0$, $\text{E}(\tilde{\delta}) \to \delta$ and $a_j\tilde{\delta} \to a_j\delta$. Second, as $\eta_j \to 0$, $\widetilde{\text{QOLI}}_j \to \text{QOLI}_j$. Third, in the bivariate example above, the sign of the amenity price bias, $\text{E}(\tilde{\delta}) - \delta$, will be determined by the product of the signs of κ, the true effect of the omitted

[10]However, even a small amount of unobservables can severely bias hedonic estimates. See e.g. Hwang, Reed, and Hubbard (1992) for a quantification of this bias in the context of wage compensating differentials.

variable, and the correlation between the included and the excluded amenity.[11] In contrast to the omitted variable bias of the implicit price, however, the size of the quality-of-life index bias, $E(\widetilde{\text{QOLI}}_j) - \text{QOLI}_j$, is inversely related to the correlation between the observed and unobserved amenity. Thus, ceteris paribus, the index bias grows as the correlation between a_j and z_j becomes more negative.[12]

Figure 3.3: Omitted variable bias in quality-of-life index estimates



Notes: Kernel density estimates for different specifications of quality-of-life indices where the indices were computed using the metropolitan sample. "Baseline QOLI" designates the quality-of-life index based on the original BBH amenities ($K = 16$). $^*\rho$ indicates correlation coefficient of the omitted variable with teacher-pupil ratio.

There are broadly three approaches for dealing with the omitted variable problem in the quality-of-life literature. First, early studies such as the work of Roback (1982) and BBH simply ignore the problem, largely because of a lack suitable amenity data. Recognising the shortcomings of this approach, Kahn (1995) highlights the difficulties that arise in the context of unobserved location-specific attributes that are correlated with observable amenities. Recent work by Gabriel and Rosenthal (2004), Shapiro (2006) and Albouy (2008) deals with unobservables by constructing quality-of-indices directly from the λ fixed-

[11] Unlike the bivariate case, anticipating the direction of the bias in the multivariate case is somewhat more complex, since it cannot be expressed as a function of the simple bivariate correlation coefficients between the omitted variable and the included variable, see e.g. Wooldridge (2006).

[12] In the bivariate case, this follows directly from $E(\widetilde{\text{QOLI}}_j) - \text{QOLI}_j = a_j(\delta + \rho_{az}\frac{\sigma_z}{\sigma_a}\kappa) - (a_j\delta + z_j\kappa) = (a_j\rho_{az}\frac{\sigma_z}{\sigma_a} - z_j)\kappa$.

effect estimates in equations (A-9) and (A-10). This approach implies, however, that there are no level effects or location-specific local shocks (i.e. $\alpha = 0$ and $\xi_j = 0$ in equation A-11). Our two-step method offers a third alternative, accounting for unobservables via fixed-effects while dealing with the index analog of the traditional omitted variable problem by systematically assembling one of the most comprehensive amenity data sets in the quality-of-life literature. The degree of the omitted variable bias on quality-of-life index estimates is illustrated in the next section.

Quantifying the omitted variable bias

Using the original BBH amenities ($K = 16$) and designating one additional amenity from our set of new amenities as "omitted variable", we provide a further illustration of the above discussion by quantifying the relative importance of the omitted variable bias for the quality-of-life indices in the sample of metropolitan counties. By iteratively adding one of several variables – namely "golf courses", "mortality", "distance to metro center", "restaurants and bars", "urban arterials", "river mileage" and "live performance venues" – we examine both the impact on the precision of specific marginal amenity price estimates and the sensitivity of the aggregate quality-of-life index to omitted variable bias. We chose the variables such that they would span a representative spectrum of correlations with the variable "teacher-pupil ratio", the proxy for school quality in the original set of BBH amenities.[13] Figure 3.3 depicts kernel density estimates of the quality-of-life index distributions for the baseline scenario of an index that includes only the amenity prices for the BBH amenities and for a set of separate indices that all include a different "omitted variable". The dependency of the index bias on the correlation of the omitted variable is clearly visible as the dispersion of the index distributions relative to the baseline index becomes more nuanced for higher, positive correlations.[14]

Table 3.1 summarises the relationship between the omitted variable bias of aggregate quality-of-life index estimates and the bias of individual amenity price estimates. For example, the baseline estimate of the marginal price of school quality as proxied by the teacher-pupil ratio, is \$4,518.57 and the corresponding aggregate quality-of-life index for the original sixteen BBH amenities has a mean of \$10,513.04. If we now compare the impact of including variables which have no or little correlation with school

[13]Our interest in variables correlated with school quality is twofold. First, the sensitivity of hedonic point estimates of the willingness to pay for school quality due to omitted variable bias is well documented (e.g. Brasington, 1999). Second, the large range of amenity price estimates is also a prominent feature of the quality-of-life literature, see table 3.5.

[14]The rank correlations of the baseline index and the indices with the additional variables confirm this relationship, ranging from 0.926 in the low correlation case to 1.000 for the high correlation case. See also table 3.1.

Table 3.1: Omitted variable bias and marginal amenity prices

		Omitted variable						
		low correlation			*high correlation*			
	Baseline	Golf course	Mortality	Metro dist.	Rest. & bars	Arterials	Rivers	Live arts
Marginal price	–	\$41.95	\$-591.58	\$-5.10	\$1,660.46	\$-373.98	\$422.45	\$1,729.14
		(5.04)	(62.27)	(0.69)	(309.66)	(790.76)	(175.71)	(1,225.42)
Teacher-pupil ratio	\$4,518.57	\$4,356.71	\$4,228.24	\$4,911.75	\$-346.95	\$5,091.03	\$1,624.14	\$-1,236.48
	(1,777.11)	(1,745.12)	(1,737.92)	(1,750.94)	(1,981.47)	(2,143.52)	(2,126.10)	(4,413.68)
	⋮	⋮			⋮			⋮
Precipitation	51.06	54.37	51.52	52.16	49.06	51.84	48.99	49.39
	(8.83)	(8.68)	(8.63)	(8.71)	(8.78)	(8.98)	(8.82)	(8.85)
Heating degree days	-1.76	-1.83	-1.53	-1.71	-1.77	-1.76	-1.68	-1.74
	(0.13)	(0.13)	(0.13)	(0.13)	(0.14)	(0.13)	(0.14)	(0.14)
Cooling degree days	-4.84	-4.97	-4.34	-4.65	-4.73	-4.84	-4.66	-4.79
	(0.21)	(0.20)	(0.21)	(0.20)	(0.21)	(0.21)	(0.21)	(0.21)
Coast	1,906.58	1,554.24	1,689.28	1,803.64	1,803.21	1,955.19	1,773.61	1,882.04
	(193.55)	(194.79)	(190.97)	(191.14)	(194.96)	(216.72)	(199.01)	(194.20)
Visibility	-167.53	-172.60	-174.13	-155.57	-164.82	-168.18	-155.93	-162.44
	(11.97)	(11.78)	(11.73)	(11.85)	(11.90)	(12.04)	(12.55)	(12.44)
City	1,648.34	1,618.94	1,633.92	1,625.33	1,527.87	1,663.00	1,464.83	1,606.40
	(260.91)	(256.22)	(255.31)	(257.12)	(259.87)	(262.03)	(264.71)	(262.24)
Correlation*	–	-0.126	-0.006	-0.001	0.466	0.537	0.551	0.887
QOLI correlation†	–	0.926	0.951	0.987	0.997	0.966	1.000	0.999
QOLI mean	\$10,513.04	\$6,122.84	\$9,402.30	\$13,530.40	\$10,245.86	\$10,666.30	\$10,514.89	\$10,300.14
QOLI st. dev.	\$2,552.19	\$2,762.25	\$2,597.47	\$2,584.46	\$2,540.48	\$2,706.58	\$2,556.39	\$2,485.27

Notes: Quality-of-life indices were computed using the metropolitan sample. "Baseline QOLI" designates the quality-of-life index based on the original BBH amenities ($K = 16$). Standard errors in parentheses. All prices are in 2000 dollars. * Sample correlation of "omitted variable" with teacher-pupil ratio. † Rank correlation of new quality-of-life index with baseline QOLI.

quality versus those which are highly correlated, two distinct features emerge. First, as we expect, including a previously omitted low-correlation variable does not have a significant impact on the estimate of the marginal willingness to pay for the teacher-pupil ratio. The inclusion of a high-correlation variable, on the other hand, yields dramatically different price estimates. Furthermore, including both low- and high-correlation variables only has comparatively muted effects on index variables other than the teacher-pupil ratio. Second, the aggregate index analog of the traditional omitted variable bias works in the opposite direction of the effect for individual prices. Omitting low-correlation variables thus has the most pronounced effect on the overall quality-of-life indices. At the same time, the omission of highly correlated variables is only of minor consequence for the aggregate index. Indeed, as is shown in the bottom of table 3.1, the aggregate index mean in the latter case is bounded within a tight range between -3% and +1% of the baseline value, whereas it fluctuates from -42% to +29% of the baseline in the former case.

Because of our interest in accurate quality-of-life index estimates rather than precise marginal prices, the collection of an expanded set of amenities plays a particular important role in our overall goal of tracking implicit household expenditure for broad categories of non-market goods and services. This is discussed next.

3.5 Data and calibration

Central contributions of this paper are the significant increase in the set of amenities upon which the quality-of-life rankings will be based and the substantial broadening of the geographical coverage for the rankings. Both aspects are discussed in turn before we describe the data for the wage and housing hedonic regression in some detail. Particular emphasis is given to improved measures of imputed rents, a further contribution that this paper makes to the hedonic quality-of-life literature.

3.5.1 Amenities

Empirical work on quality-of-life rankings is frequently limited by the availability of suitable amenity data at a sufficiently high level of geographic resolution. Beyond simple data considerations, work on amenity valuation might also be guided by the researcher's priors on what types of amenities people value the most. Usually, it is assumed that – in addition to popular local public goods such as the school quality and the absence of crime – households care the most about high environmental quality, such as clean air, and popular geographical features and nice weather. Indeed, of the total of sixteen amenities that are considered in BBH, nine are what we classify as geographical amenities (which predominantly include climate variables) and another six are environmental amenities. As a direct consequence of this relatively narrow focus on a specific type of amenity, visual representations of such quality-of-life rankings (or predicted growth patterns based thereon) often simply resemble maps of weather patterns or climate zones.[15]

Table 3.2: Amenity summary statistics – BBH urban counties

	1980 – BBH		2000			
	Mean	*Mean*	*Std. Dev.*	*Min.*	*Max.*	*Sources**
GEOGRAPHY AND CLIMATE						
Mean precipitation (inches p.a., 1971–2000)	32.00	38.13	13.54	9.81	101.96	NOAA-NCDC
Mean relative annual humidity (%, 1961–1990)	68.30	67.64	7.40	30.50	78.00	NOAA-NCDC
Mean annual heating degree days	4,326.00	4,653.45	2,051.88	214.49	9,608.38	NOAA-NCDC
Mean annual cooling degree days	1,162.00	1,289.17	851.49	105.33	3,966.34	NOAA-NCDC
Mean wind speed (m.p.h., 1961–1990)	8.89	8.91	1.07	6.41	11.55	NOAA-NCDC
Sunshine (% of possible)	61.10	59.51	8.04	45.91	82.72	NOAA-NCDC
Heavy fog (no. of days with visibility ≤ 0.25 mi.)	15.80†	20.20	8.10	2.70	45.25	NOAA-NCDC
Percent water area	–	9.94	15.94	0.03	75	ICPSR
Coast (=1 if on coast)	0.33	0.24	0.65	0	1	NOAA-SEAD
Non-adjacent coastal watershed (=1 if in watershed)	–	0.13	0.33	0	1	NOAA-SEAD
Mountain peaks above 1500 meters	–	7.10	26.04	0	191	ESRI
Rivers (miles per sq. mile)	–	0.24	0.14	0.027	1.34	USDI-NPS
Federal land (percentage of total land area)	–	9.17	19.78	0	100	USGS-NA
Wilderness areas (percentage of total land area)	–	1.14	3.86	0	23.44	USGS-NA
National Parks (percentage of total land area)	–	0.8	2.75	0	26.88	USGS-NA
Distance (km) to nearest National Park	–	71.81	54.74	3.22	303.19	USDI-NPS
Distance (km) to nearest State Park	–	22.68	17.24	0.48	150.78	USDI-NPS
Scenic drives (total mileage)	–	0.21	0.8	0	10.00	USGS-NA
Average number of tornados per annum (1950–2004)	–	0.44	0.54	0	4.13	USGS-NA
Property damage from hazard events ($000s, per mi^2)	–	57.86	128.26	0.04	1,502.95	USGS-NA
Seismic hazard (index)	–	2,029.3	327.96	1,022.50	3,125.90	USGS-NA
Number of earthquakes (1950–2000)	–	3.36	16.91	0	178	USGS-NA
Land cover diversity (index, range 0–255)	–	146.37	37.37	35.38	217.84	USGS-NA
ENVIRONMENTAL EXTERNALITIES						
NPDES effluent dischargers (PCS permits, 1989–1999)	1.51	16.67	32.51	0	209	EPA-TRI
Landfill waste (metric tons, 2000)	4,770.00	4,106.13	25,474.37	0	351,877.40	EPA-TRI
Superfund sites	0.88	2.73	3.71	0.00	23.00	EPA-TRI
Treatment, storage and disposal facilities	46.40	34.42	59.80	0	570	EPA-TRI
Large quantity generators of hazardous waste	–	218.45	359.82	0	3,652	EPA-TRI
Nuclear power plants	–	0.06	0.25	0.00	2.00	USDOE-INSC
PM2.5 (μg per m^3)	–	13.56	2.92	7.14	22.51	EPA-AQS

continued on the next page

[15]Such patterns can of course be intentional as in Rappaport (2007) where a large portion of the migration patterns appears to be driven by an increased valuation of nice weather as a consumption amenity. Glaeser and Tobio (2008), however, do not find any evidence of an increase in the willingness to pay for sun-related amenities, but rather attribute growth of the Sun Belt to increases in the housing supply.

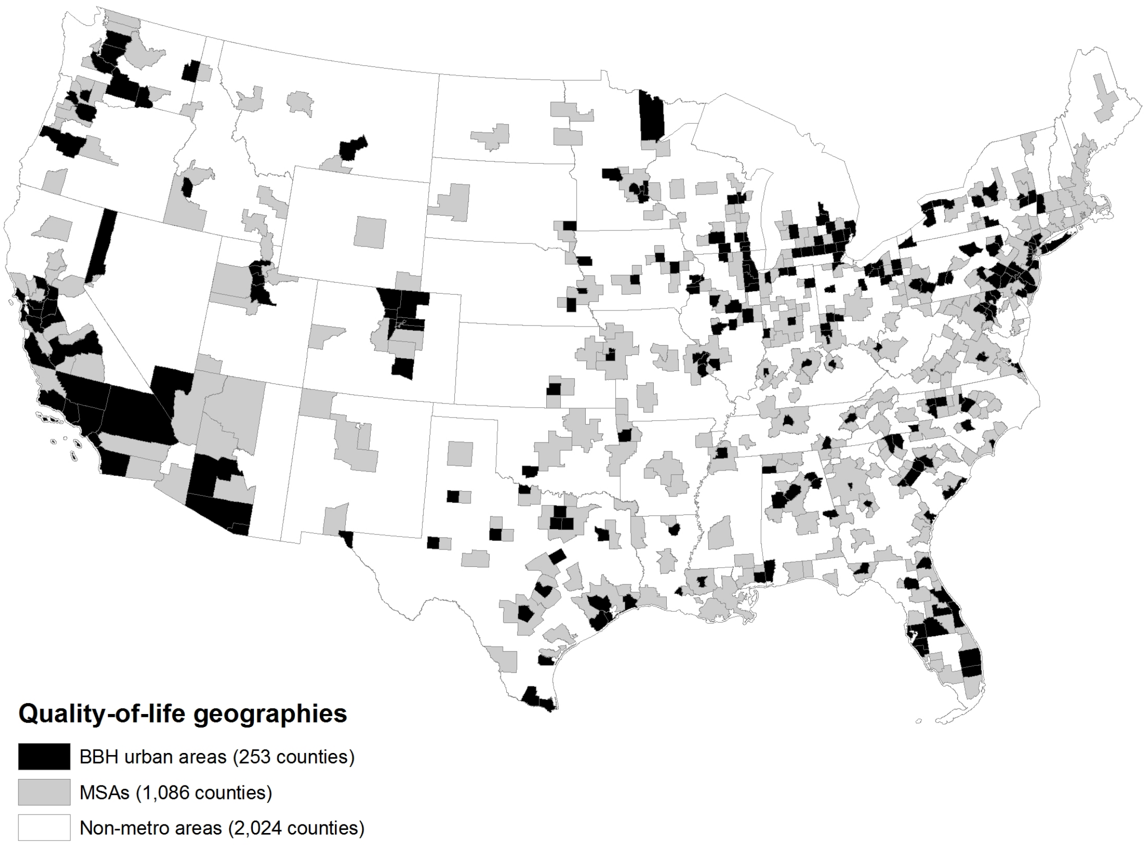
	1980 – BBH	2000				
	Mean	*Mean*	*Std. Dev.*	*Min.*	*Max.*	*Sources**
PM10 (μg per m^3)	73.20‡	23.73	5.30	5.00	47.05	EPA-AQS
Ozone (μg per m^3)	–	9.78	12.61	0.02	39.99	EPA-AQS
Sulphur dioxide (μg per m^3)	–	1.46	1.87	0.00	9.83	EPA-AQS
Carbon monoxide (μg per m^3)	–	6.04	22.85	0.07	249.67	EPA-AQS
Nitrogen dioxide (μg per m^3)	–	5.59	5.84	0.00	25.05	EPA-AQS
National Fire Plan treatment (percentage of total area)	–	0.11	0.48	0	4.69	USGS-NA
Cancer Risk	–	4.08	1.67	1.17	13.61	EPA-NATA
Neurological risk	–	0.10	0.06	0.03	0.43	EPA-NATA
Respiratory risk	–	5.31	3.66	0.86	29.40	EPA-NATA
LOCAL PUBLIC GOODS						
Local direct general expenditures ($ per capita)	–	3.44	0.95	1.83	10.80	COG97
Local exp. for hospitals and health ($ per capita)	–	49.2	97.35	0	793.72	COG97
Local exp. on parks, rec. and nat. resources ($ pc)	–	15.4	33.25	0	351.24	COG97
Museums and historical sites (per 1,000 people)	–	0.04	0.08	0	1.79	CBP
Municipal parks (percentage of total land area)	–	1.54	2.31	0	13.87	ESRI
Campgrounds and camps	–	6.31	7.72	0	54	CBP
Zoos, botanical gardens and nature parks	–	1.77	2.39	0	19	CBP
Crime rate (per 100,000 persons)	647	4,692.25	6,030.59	139	96,058	ICPSR
Teacher-pupil ratio	0.080	0.091	0.058	0.038	0.665	COG97
Local expenditure per student ($, 1996-97 fiscal year)	–	37.31	124.59	0.03	1257.88	COG97
Private school to public school enrollment (%)	–	13.41	10.32	4.02	64.46	2000 Census
Child mortality (per 1000 births, 1990–2000)	–	6.99	1.46	3.99	12.32	CDC-NCHS
INFRASTRUCTURE						
Federal expenditure ($ pc, non-wage, non-defense)	–	4,995.72	6,450.05	0	131,029.30	COG97
Number of airports	–	2.12	2.2	0	15	USGS-NA
Number of ports	–	0.27	0.58	0	4	USGS-NA
Interstate highways (total mileage per mi^2)	–	0.09	0.08	0	0.56	USGS-NA
Urban arterial (total milage per mi^2)	–	0.25	0.45	0	5.53	USGS-NA
Number of Amtrak stations	–	1.16	1.65	0	15	USGS-NA
Number of urban rail stops	–	7.38	27.4	0	317	USGS-NA
Railways (total mileage per mi^2)	–	0.45	0.53	0	5.85	USGS-NA
CULTURAL AND URBAN AMENITIES						
Number of restaurants and bars (per 1,000 people)	–	1.01	0.78	0	19.71	CBP
Theatres and musicals (per 1,000 people)	–	0.01	0.04	0	1.20	CBP
Artists (per 1,000 people)	–	0.18	0.34	0	3.94	CBP
Movie theatres (per 1,000 people)	–	0.02	0.05	0	1.27	CBP
Bowling alleys (per 1,000 people)	–	0.03	0.06	0	0.82	CBP
Amusement, recreation establishments (per 1,000 people)	–	0.07	0.19	0	4.81	CBP
Research I universities (Carnegie classification)	–	0.23	0.53	0	3	CCIHE
Golf courses and country clubs	–	15.94	15.43	0	123	CBP
Military areas (percentage of total land area)	–	1.19	2.89	0	19.3	USGS-NA
Housing stress (=1 if > 30% of hholds distressed)	–	0.58	0.49	0	1	USDA-ERS
Persistent poverty (=1 if > 20% of pop. in poverty)	–	0.02	0.13	0	1	USDA-ERS
Retirement destination (=1 if growth retirees > 15%)	–	0.08	0.26	0	1	USDA-ERS
Distance (km) to the nearest urban center	–	11.51	13.43	0	68.63	PRAO-JIE09
Incr. distance to a metropolitan area of any size	–	0.66	6.27	0	89.10	PRAO-JIE09
Incr. distance to a metro area > 250,000	–	23.91	67.04	0	584.49	PRAO-JIE09
Incr. distance to a metro area > 500,000	–	33.26	64.67	0	479.20	PRAO-JIE09
Incr. distance to a metro area > 1.5 million	–	77.95	125.84	0	599.21	PRAO-JIE09

Notes: * The amenity data is constructed from base data from the following sources: CCIHE: Carnegie Classification of Institutions of Higher Education; CBP: 2000 County Business Patterns published by the Census Bureau; CDC-NCHS: Centers for Disease Control and Prevention, National Center for Health Statistics; COG97: 1997 Census of Governments; EPA-AQS: 2000 data for criteria air pollutants from the Air Quality System produced by the Environmental Protection Agency (EPA); EPA-NATA: 1999 National-Scale Air Toxics Assessment conducted by the EPA; EPA-TRI: 2000 Toxic Release Inventory published by the EPA; ESRI: Environmental Systems Research Institute ArcGIS maps; ICPSR: U.S. County characteristics complied by the Inter-university Consortium for Political and Social Research (ICPSR, 2008); NOAA-SEAD: Strategic Environmental Assessments Division of the National Oceanic and Atmospheric Administration; NOAA-NCDC: National Climatic Data Center of the National Oceanic and Atmospheric Administration; PRAO-JIE09: Partridge, Rickman, Ali, and Olfert (2009); USDA-ERS: Economic Research Service of the US Department of Agriculture; USDI-NPS: National Park Service of the US Department of the Interior; USDOE-EERE: Energy Efficiency and Renewable Energy, US Department of Energy; USDOE-INSC: International Nuclear Safety Center at the US Department of Energy; USGS-NA: National Atlas of the US Geological Survey. † The unit in the BBH visibility variable is miles, rather than total days with a minimum visibility of less than 0.25 miles. ‡ BBH use data on total suspended particulates (TSP), a precursor measure to PM10.

Overall we collect over 60 different amenities of which sixteen are geographical amenities, sixteen are environmental amenities, twelve are neighbourhood amenities, eight are infrastructure amenities and sixteen are urbanisation amenities. Table 3.2 lists all of our amenity variables, grouped by these five categories. The table also compares the means of the BBH amenities in 1980 to their updated values in 2000. We discuss specific changes in the levels of these amenities over the last 20 years, their possible causes and – most importantly their impact on the rankings – in section 3.7. Specific methodological issues related to the amenity data is discussed in section 3.B of the appendix.

While expanding the set of geographical and environmental amenities, we also try to extend our focus to a significant collection of neighbourhood amenities, infrastructure amenities and urbanisation amenities. Our expanded set of neighbourhood amenities is largely motivated by Gyourko and Tracy (1991) who show that differences in local fiscal conditions generate compensating differentials where by the fiscal climate affects the quality-of-life. The largest of our additional categories are urbanisation or consumption amenities. Recent work on the "consumer city" assigns an increasing importance on this type of amenity, both as a driver of urban growth, agglomeration and in the determination of urban wage structures (Glaeser, Kolko, and Saiz, 2001; Glaeser and Gottlieb, 2006; Krupka, 2008; Lee, 2010). Similarly, infrastructure amenities are also likely to play an important role in the determination of the quality-of-life. This effect might be particularly prominent in urban areas, where the link between (transportation) infrastructure and congestion disamenities is well-documented. In addition to this direct channel, there are more subtle ways in which infrastructure amenities might influence the quality-of-life via their impact on the built environment. For example, public infrastructure has a direct influence on the spatial characteristics of urban structure (Burchfield, Overman, Puga, and Turner, 2006; Baum-Snow, 2007a,b). The specific mix of infrastructure amenities in a given location is also the result of a combination of unique local policy outcomes, such as anti-sprawl measures or urban renewal. In this sense, infrastructure amenities additionally capture local residents' public policy preferences and their consequences.[16]

Figure 3.4: Variation of geographical coverage



Notes: BBH denotes the urban counties originally included in Blomquist, Berger, and Hoehn (1988), using data from the 1980 Census. In 1980, these counties covered 49% of the US population, while only accounting for 9.3% of the land area of the contiguous US. By the 2000 Census the BBH counties corresponded to 47% of the total population. The 1,086 counties that belong to a metropolitan area in 2000 cover 80.3% of the total population while covering 29.7% of the land area. See also table 3.3 for a more detailed description of the different samples.

[16] See Anas and Pines (2008) for a formal model that links anti-sprawl policies and cities that differ in amenities. Indirectly, a particular combination of fiscal policies by local government co-determines the spatial pattern both public and private infrastructure which in turn influences urban structure. Recent evidence on the link between specific tax regimes and urban structure is provided by Banzhaf and Lavery (2010) and Song and Zenou (2009).

3.5.2 Geographic coverage

On top of increasing the number of amenities, this paper also considers a substantial increase in the geographical scope while maintaining a relatively high level of spatial resolution. Previous work with a comparable geographical reach is limited to analysis at the state-level for even smaller set of amenities than BBH (e.g. Gabriel, Mattey, and Wascher, 2003).[17] Overall, we evaluate quality-of-life estimates for three spatial samples which are illustrated in figure 3.4. Our first sample consists of the 253 counties that were originally covered in BBH. While only representing 9.3% of the total land area, as the total US population grew from 226.5 million in 1980 to 279.6 million in 2000, the population density in these urban areas increased by almost 20% from 419 to 500 people per sq.mi., compared to the national average of 77 and 94 people per sq. mi respectively. At the same time, however, the total population coverage of these high-density counties fell from 48.8% in 1980 to 47.2% in 2000 of the total US population, indicating that other counties experienced more population growth, most likely in response to rising urban congestion. This trend in the urbanisation process helps us define our second geographic sample which consists of the 1,086 counties that were part of a metropolitan statistical area (MSA) following the Office of Management and Budget definition used in the 2000 Census. These counties increase our population coverage up to 80.3% of the total population and represent 29.7% of the surface area of the contiguous US. The rapid pace of US urban development is particulary salient when considering that the population density in these counties rose from 133 to 259 people per sq. mi. in just two decades.[18]

Because public use microdata areas (PUMA) that receive the 5-percent data must have a minimum census population of 100,000, the geographic area of PUMAs varies inversely with population density, averaging around 259 people per square mile at the county-level. In most urban areas, there are several PUMAs that fall within a given county, whereas a PUMA can span several counties in the most sparsely populated areas. For example, the most densely populated county (New York County, NY) has a population density of 66,951 people per sq.mi and is covered by ten PUMAs. By contrast, in Loving County, TX, which is both the least populous and the least densely populated county in the US, there are only 0.09 people per sq.mi. and its corresponding PUMA covers fourteen counties. This PUMA-to-county

[17] In contrast to Gabriel, Mattey, and Wascher (2003), however, we do not systematically consider changes in the quality-of-life rankings over time.

[18] In 1980, only 698 of the 1,086 counties that were part of an MSA in 2000 were defined as metropolitan according to the 1980 definition which covered a total 730 counties. The 32 counties that were considered metropolitan in 1980, but were no longer classified as part of an MSA in 2000 fall mostly into the newly created category of mircopolitan statistical areas (μSA) which are agglomerations with a population below 50,000.

Table 3.3: Characteristics of geographical coverage

		Geography		
		BBH urban counties	*Metropolitan counties**	*All counties*†
No. of counties		253	1,085	3,110
No. of PUMAs		1,061	1,835	2,057
PUMAs per county		4.19	1.69	0.67
Population	*1980*	110,617,710	170,867,817	226,545,805
	2000	131,860,476	224,482,276	279,583,437
Pop. coverage	*1980*	48.8%	75.4%	100.0%
	2000	47.2%	80.3%	
Pop. density (inh. per mi^2)	*1980*	419	402	77
	2000	500	259	94
Land area (mi^2)		263,840	865,437	2,959,064
Water area (mi^2)		25,273	61,081	160,820
Total area (mi^2)		289,113	926,518	3,119,885
Areal coverage		9.3%	29.7%	100.0%
N. obs from PUMS		4,833,916 (P)	8,875,172 (P)	10,198,936 (P)
		2,587,457 (H)	4,795,515 (H)	5,484,870 (H)

Notes: Public use microdata areas (PUMA) that receive the 5-percent data must have a minimum census population of 100,000. * 1980 or 2000 OMB definitions of metropolitan statistical areas (MSA) where applicable. † Contiguous U.S. only. *Source*: Authors' calculations using Census data.

relationship is summarised in table 3.3. Since virtually all amenity data is only available at the county-level, we assign each PUMA its amenity values by applying the appropriate population weights to the corresponding county-level amenity data.[19]

3.5.3 Data on wages and housing expenditures

The data on wages and housing expenditure were obtained from the 5% public-use microdata sample (PUMS) of the 2000 Census. In line with the literature, we construct hourly wages from reported annual earnings as the dependent variable for the wage regressions and we use monthly housing expenditure imputed from self-reported housing values and rents for the housing hedonic regression. Deriving both hourly wages and a measure of housing expenditure from PUMS data represents a number of technical and theoretical challenges that are discussed in more detail in sections 3.A.1 and 3.A.3 of the appendix. Nonetheless, some of the conceptional issues related to evaluating the most suitable measure for monthly housing expenditure warrant a more detailed discussion here.[20]

[19] A more detailed account of our methodology for dealing with county-level amenity data and PUMAs is discussed in section 3.B of the appendix.

[20] The more technical issues related to using data on housing values that are self-reported and grouped across intervals are addressed in section 3.A.3.

The importance of imputed rents

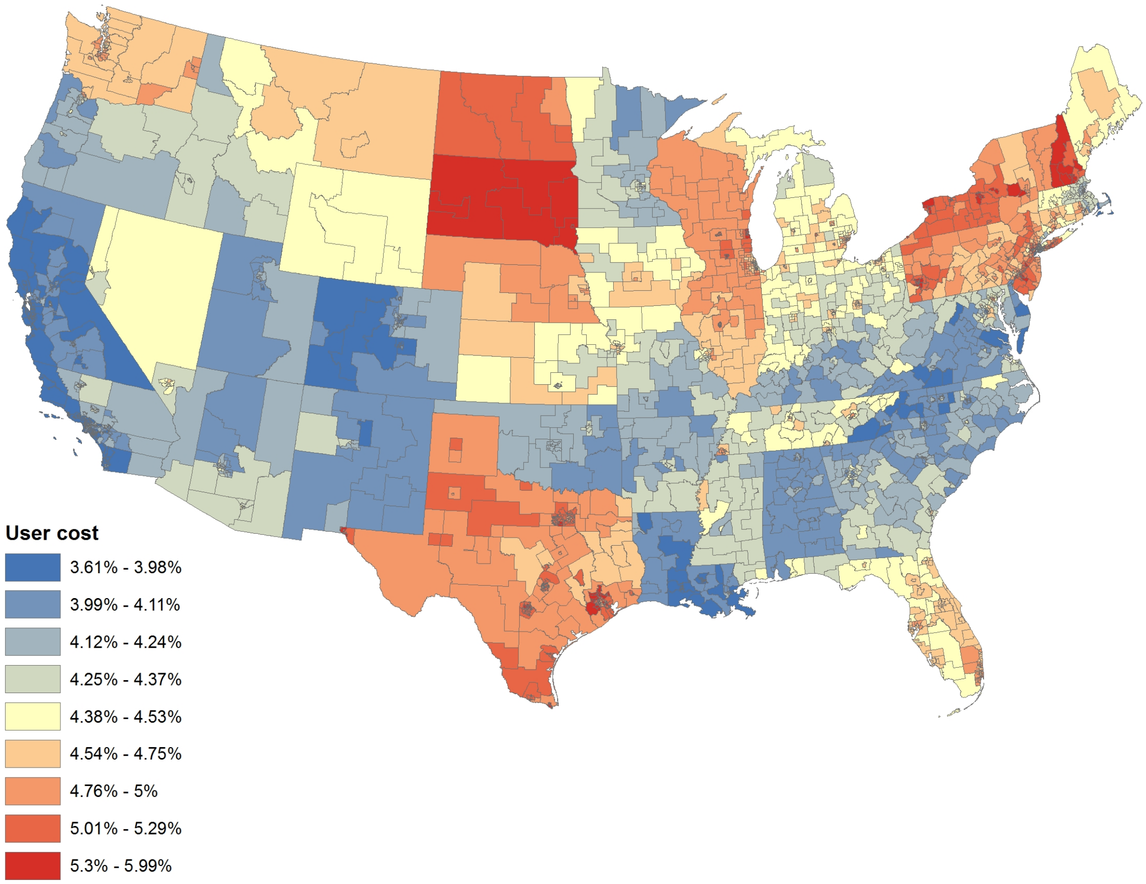
Given that the US homeownership rate is around 67.5%, the measurement of housing expenditures for owner-occupier households – as opposed to renters – plays a pivotal role in determining location-specific quality-of-life estimates. In 2003, some 40 million households each claimed an average of $9,500 in mortgage interest deduction and almost $3,000 in property tax deductions. This renders the subsidy to homeowners one of the most prominent features of the American tax code. Moreover, the incident of these tax-related benefits to homeownership is very unevenly distributed across space (Gyourko and Sinai, 2003). Against this backdrop, appropriately accounting for the local differences in housing expenditure seems at least as important as other methodological refinements of quality-of-life estimates.

In contrast to the bulk of existing work, we do not assume a fixed, national discount rate by which housing values are converted into implied housing expenditures. Instead, we adopt the approach of Himmelberg, Mayer, and Sinai (2005) to construct a more appropriate measure for each owner-occupied household, the user cost of housing or "imputed rent", which we denote by ϕ. BBH use a fixed discount rate of 7.86% which is based on simulations results by Peiser and Smith (1985, p.357) for an ownership interval from 1987-1990 under a scenario of anticipated rising inflation. Despite the secular changes in the interest rate environment that lead to unprecedented house price increases over last three decades, more recent studies have paid very little attention to this point and use that same constant rate (e.g. Gyourko and Tracy, 1991; Gabriel and Rosenthal, 2004; Albouy, 2008; Chen and Rosenthal, 2008). This cost of homeownership then reflects the most accurate measure of the opportunity cost of owning a house that is directly comparable to rental costs. As we discuss in section 3.7.2, our improved measurement of the user cost of housing has a much more significant impact on quality-of-life estimates that other improvements, such as using after-tax wages or using a broader measure of the cost-of-living that goes beyond housing cost alone.

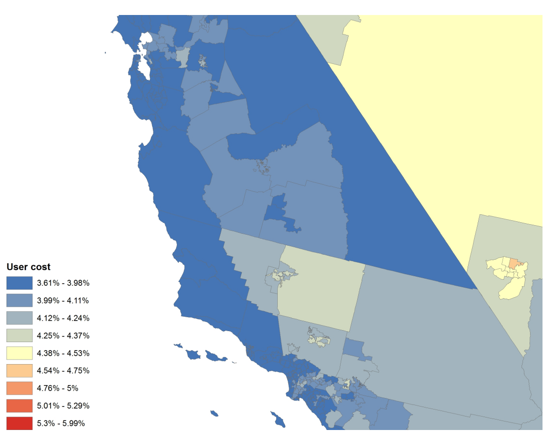
Spatial variations in the user cost of housing

The large local differences in the user cost of housing across the US are illustrated in figure 3.5. Imputed rents do not only vary greatly between regions, but there are also significant intra-urban differences. Our sample yields a national average for user cost of $\bar{\phi} = 5.12\%$, with a minimum of 4.16% and a maximum value of 9.89%. This implies a national range for the price-to-rent ratios of 24.0 to 10.1 with an average ratio of 19.5, suggesting that on average people should be willing to pay up to twenty times the market

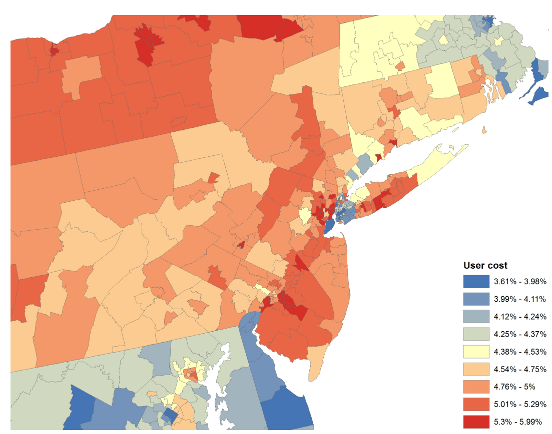
Figure 3.5: Spatial distribution of user cost across PUMAs



(a)



(b) Central California



(c) East Coast, DC–NYC–Boston

Notes: User cost of housing are the discount factor by which imputed rents are calculated from self-reported house values. They reflect several components that influence house ownership, namely the opportunity cost of foregone alternative investments, property taxes, tax subsidies, maintenance cost, expected capital gains and a risk premium. See appendix 3.A.3 above for a more detailed description of the methodology. *Source*: Author's calculations.

rent to purchase a house. Compared to the national sample, the average user cost for the BBH sample of urban counties is only marginally higher at $\phi^{\text{BBH}} = 5.16\%$.[21]

Our calculations introduce two sources of location-specific variation for the user cost: local differences in property taxes and differences in average state and federal income taxes. Differences in property tax rates across municipalities do not only arise endogenously as part of the Tiebout sorting process, but there are also state-level effects due to specific legislation that limits the level of property taxes, such as e.g. Proposition 13 in California (1978), Proposition $2\frac{1}{2}$ in Massachusetts (1980), or Ballot Measure 5 in Oregon (1990). While the burden of federal taxes on households in high-wage urban areas is higher compared to otherwise identical households who live in low-wage urban areas (Albouy, 2009a), the overall economic impact of this burden can be partially offset by mortgage and local-tax deductions that disproportionately benefit high-cost areas.[22]

Assuming a uniform user cost of housing does not take these effects into consideration and might thus distort the quality-of-life rankings. Since imputed rent in location j is defined as $r^j_{\text{imp}} = P^j \phi^j$, using a uniform user cost rate $\bar{\phi}$ instead of letting ϕ^j vary across locations, ceteris paribus, tends to overstate rents in low cost regions and to underestimate rents in high cost regions. As a result, regions where user costs of housing are below the national average appear to have a higher quality-of-life than they actually do in reality and regions with higher user cost appear to have a lower quality-of-life. This distortion in equilibrium rent differentials implies a potential bias in the quality-of-life index ranking by overranking low cost localities on the West coast and underranking higher cost areas in Texas, parts of the Mid-West and the North-East. We quantify the sensitivity of the rankings to the uniform user cost assumption in our discussion of the empirical results in section 3.7.2.

3.6 Results from hedonic estimation

The results of the hedonic housing and wage regression from using a single-stage BBH-style estimation approach and the first stage represented by equations (A-9) and (A-10) of our preferred, two-stage specification are summarised in table 3.4. As a benchmark for comparison, we reproduce the one-stage results of BBH as model (1) in the first two columns of the table. The BBH empirical set-up for both the housing and wage regression involves a linear Box-Cox transformation of the dependent variable (log monthly

[21] By comparison, the fixed rate discount factor of 7.86% used in BBH implies a significantly lower price-to-rent ratio of 12.7.

[22] See section 3.A.3 for a more detailed discussion of this point.

housing expenditure and log hourly wages) using a separate transformation parameter θ for each of the hedonic regressions. We replicate the BBH set-up with our updated data set in two slightly modified ways. First, we apply the same type of Box-Cox transformation to the dependent variable only in model (2) and then we estimate a different Box-Cox version in model (3) where the independent variables are also transformed using a different parameter λ. Our models (4) and (5) are estimated using the two-stage approach, where model (4) is limited to the amenities originally included in BBH and model (5) includes all of our additional amenities presented in the previous section.

Most parameter estimates are significant across all four specifications with the common result that the respective fit for the housing regression tends to be substantially higher than the fit of the hedonic wage equation. Both Box-Cox specifications yield largely similar estimates although there are significant differences in terms of the level and sign – particularly for the wage regressions – compared to estimates in BBH. The sensitivity of the optimal Box-Cox transformation parameters to the underlying data set is also illustrated by the span of values for θ and λ when comparing model (1) to models (2) and (3). While the parameters estimates from the two-stage estimation in models (4) and (5) cannot be directly compared to those from the one-stage estimation, the first stage fit of these model (not reported here) is similar. More importantly, the fit of the second-stage model improves significantly by adding the new set of amenities in model (5). The location-specific variation in amenities explains over 70% of the variation in the rent differentials and close to 50% of the variation in the compensating differential as measured by the area fixed-effects from the first-stage housing and wage regressions. The two-stage set-up of model (5) – corresponding to equations (A-9) and (A-12) – thus provides our benchmark estimates for the remainder of our analysis. Overall, while the sensitivity of parameter estimates to the precise econometric specification of the hedonic model is well-documented (Cropper, Deck, and McConnell, 1988; Kuminoff, Parmeter, and Pope, 2010), its implications for quality-of-life rankings are best illustrated by the corresponding variation in the implicit amenity prices which are derived from these estimates.

3.6.1 Marginal amenity price estimates

In table 3.5, we derive the full implicit amenity prices that correspond to the different specifications of the housing and wage regressions which were presented in table 3.4. These prices represent the average household's marginal valuation for a given amenity as implied by the average rent and wage differential that prevail in specific location. As a point of departure, column (1) lists the current-dollar equivalent of the full implicit prices for the sixteen amenities originally reported in BBH. In column (2), we add

Table 3.4: Hedonic estimates of housing and wage differentials

	BBH		*BBH update*†				*2-stage estimation*†			
	Box-Cox (LHS)		Box-Cox (LHS)		Box-Cox (full)		FGLS		FGLS	
	(1)		(2)		(3)		(4)		(5)	
	Housing‡	Wage◇	Housing	Wage	Housing	Wage	Housing	Wage	Housing	Wage
Precipitation	-1.047	-0.0144	2.9628	0.1232	4.2761	0.0328	0.5651	0.2118	-0.0675	-0.1155
	(0.1490)	(0.004)	(0.1842)	(0.0254)	(0.2760)	(0.0060)	(0.0955)	(0.0436)	(0.1297)	(0.0650)
Humidity	-2.1270	0.0065	-1.5323	0.0598	-2.6959	0.0176	-0.1642	0.0275	0.2469	-0.0038
	(0.2510)	(0.0060)	(0.3333)	(0.0455)	(0.5134)	(0.0101)	(0.1726)	(0.0787)	(0.2374)	(0.1191)
Heating degree days	-0.0136	-0.0001	-0.0988	-0.0032	-0.2606	-0.0001	-0.0106	-0.0026	0.0010	0.0016
	(0.0010)	(0.0000)	(0.0030)	(0.0004)	(0.0076)	(0.0000)	(0.0016)	(0.0007)	(0.0026)	(0.0013)
Cooling degree days	-0.0760	-0.0002	-0.2745	-0.0094	-0.6510	-0.0005	-0.0351	-0.0096	-0.0110	0.0015
	(0.0020)	(0.0001)	(0.0043)	(0.0006)	(0.0102)	(0.0000)	(0.0022)	(0.0010)	(0.0054)	(0.0027)
Wind speed	11.8800	0.0961	59.1904	2.4012	78.6574	0.9462	8.1993	2.4551	-1.9221	-1.5780
	(0.8670)	(0.0220)	(2.442)	(0.3351)	(3.1213)	(0.1445)	(1.2986)	(0.5924)	(1.9302)	(0.9680)
Sunshine	2.1350	-0.0091	5.4906	0.1260	7.5700	0.0398	1.1711	0.2370	0.6202	0.0244
	(0.2350)	(0.0060)	(0.5318)	(0.0733)	(0.8378)	(0.0146)	(0.2788)	(0.1272)	(0.3589)	(0.1800)
Coast	32.5100	-0.0310	132.9104	3.7600	120.2076	4.4647	21.0418	4.0945	14.1538	3.0521
	(2.4700)	(0.0630)	(4.0598)	(0.5609)	(4.1365)	(0.5463)	(2.1569)	(0.9839)	(4.3371)	(2.1752)
Violent crime	0.0434	0.0006	0.0021	0.0001	0.0065	0.0000	0.0002	0.0001	0.0002	0.0000
	(0.0030)	(0.0001)	(0.0001)	(0.0000)	(0.0004)	(0.0000)	(0.0001)	(0.0000)	(0.0002)	(0.0001)
Teacher-pupil ratio	635.3000	-5.4510	422.8988	33.4338	349.5911	81.6650	26.8436	14.8912	109.0503	59.0279
	(71.600)	(1.8500)	(57.0138)	(8.3746)	(46.3928)	(16.4155)	(15.8217)	(7.2234)	(23.7065)	(23.6117)
Visibility	-0.8302	-0.0026	-8.3184	-0.4872	-11.4936	-0.1458	-1.1363	-0.4586	-0.1185	-0.0955
	(0.110)	(0.0030)	(0.2859)	(0.0400)	(0.3939)	(0.0127)	(0.1566)	(0.0714)	(0.1901)	(0.0953)
TSP/PM10	-0.5344	-0.0024	3.5142	0.2294	5.8285	0.0588	0.2032	0.2269	0.0441	-0.0027
	(0.0580)	(0.0010)	(0.3398)	(0.0467)	(0.4868)	(0.0134)	(0.1787)	(0.0815)	(0.2140)	(0.1073)
NPDES effluent dischargers	-7.4580	-0.0051	-0.2143	0.0000	-0.2747	0.0022	-0.0421	-0.0075	-0.0329	-0.0171
	(0.4610)	(0.0120)	(0.0588)	(0.0082)	(0.0587)	(0.0082)	(0.0311)	(0.0142)	(0.0323)	(0.0162)
Landfill waste	0.0095	0.0001	-0.0005	0.0000	-0.0017	0.0000	-0.0001	0.0000	0.0000	0.0000
	(0.0010)	(0.0000)	(0.0001)	(0.0000)	(0.0002)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Superfund sites	13.4200	0.1069	14.9581	0.6167	14.9426	0.6124	2.4244	0.9134	0.3428	0.4433
	(0.6930)	(0.0170)	(0.4322)	(0.0582)	(0.4328)	(0.0576)	(0.2245)	(0.1024)	(0.3030)	(0.1519)
Treatment, storage & disposal sites	0.2184	0.0013	-0.3457	-0.0068	-0.3296	-0.0079	-0.0584	-0.018	-0.0096	-0.0083
	(0.0240)	(0.0010)	(0.0181)	(0.0025)	(0.0180)	(0.0025)	(0.0093)	(0.0043)	(0.0189)	(0.0095)
Central city	40.7500	-0.4537	32.5407	-3.0238	29.5392	-2.6897	4.811	-5.7826	5.8745	-6.8003
	(2.5400)	(0.0650)	(10.4068)	(1.4107)	(10.3882)	(1.4112)	(5.9439)	(2.7114)	(5.8115)	(2.9146)
New amenities	N		N		N		N		Y	
Adj. R^2	0.6624	0.3138	0.4312	0.2841	0.4325	0.2867	0.5186	0.3515	0.7148	0.4880
N.Obs.	34,414	46,004	2,395,116	3,223,602	2,395,116	3,223,602	2,395,116	3,223,602	2,395,116	3,223,602
Log-likelihood	-219,013	-124,403	-38,694	-61,796	-38,671	-61,689	–	–	–	–
Box-Cox parameter θ	0.200	0.100	0.877	0.873	0.907	0.878	–	–	–	–
Box-Cox parameter λ	1.000	1.000	1.000	1.000	2.022	1.376	–	–	–	–

Notes: Standard errors are shown in parentheses below the parameter estimates. All estimates are significant at the 5%-level except those marked by * and the coefficients for the Box-Cox regressions are linearised. † The parameter estimates are multiplied by a factor of 10^2 to facilitate readability and comparison. ‡ The dependent variable for the housing regression is actual or impute monthly housing expenditure. The following structural control variables are included in the housing hedonic equation, but are not reported: Rooms, bedrooms, size and age of building, acreage, type of unit and condominium status, quality of kitchen and plumbing facilities, renter status and renter status interaction terms for each of these variables. ◇ The dependent variable for the wage hedonic regression is hourly wages with following control variables that are not reported here: experience (age-schooling-6), experience squared, gender interaction with experience and experience squared, marital status, race, gender interaction with marital status, age and children under 18, educational attainment and/or enrollment, citizen status, employment disability, NAICS-based industry and occupational class, and military status.

the amenity prices that were reported in a study by Stover and Leven (1992, SL henceforth) who use the original BHH data set, but employ a different specification for the hedonic estimation which yields some striking differences with regard to both the level and the sign of the implicit prices. As such, the SL prices provide a second basis of comparison for our estimates. Columns (3) and (4) report the corresponding amenity prices from the 2000 data, but using the same Box-Cox estimation framework as in BBH. Columns (5) and (6) report the amenity price estimates from our 2-stage estimation, including either the original set of amenities in column (4) which is then extended to the full set of amenities in column (5).

Overall, our implicit prices display patterns similar to those originally reported by BBH, but there are some interesting divergences which are likely to originate from three separate sources. First, different empirical specifications might give rise to different price estimates for amenities. Indeed, the variation across different valuation estimates reported in table 3.4 confirm that the implicit amenity prices are subject to considerable sensitivity with regard to the econometric specification of the hedonic function. The second source of variation stems from the subjectivity that is reflected in any given selection of amenity variables, where omitted variables might bias the price estimates for some of the amenities. Lastly, the change in the estimates for the implicit prices can be purely data-driven if there is a large enough, qualitative change in the level of the amenity data over the two decades between 1980 and 2000.[23]

3.6.2 Hedonic specification

This sensitivity of the estimates to econometric specification are best illustrated by considering the case of the implicit price for one amenity in particular, namely that of school quality. Perhaps the most striking aspect of the BBH estimates is the large magnitude of the implicit price of school quality, as proxied by the teacher-pupil ratio. By comparison, while we obtain values that are of similar relative magnitude for our Box-Cox specifications, our overall range for the full implicit price of school quality is considerably lower ranging between \$1,975 in column (5) to \$34,934 in column (3) which is replicates BBH most closely. While Brasington (1999) shows that the teacher-pupil ratio is consistently capitalised into housing prices, several authors have shown that the hedonic valuation of school quality is affected by supply conditions, such as the presence of private schools (e.g. Downes and Greenstein, 1996; Downes and Zabel, 2002). In particular, the extent to which the variation in public school input levels, such as

[23] The possibility of changes in the sorting equilibrium in the housing or labour market would have to be evaluated within a fully specified, structural sorting model and are not considered there as this is beyond the scope of the Rosen-Roback approach (cf. Bayer and Timmins, 2007).

Table 3.5: Full implicit prices from housing and wage differentials

	*BBH**	*SL**	*BBH update*		*2-stage estimation*	
	Box-Cox	Box-Cox	Box-Cox	Box-Cox	FGLS	FGLS
	(1)	(2)	LHS (3)	full (4)	(5)	(6)
Precipitation	$49.11	$-28.65	$247.91	$350.63	$43.2	$-11.12
	(20.29)	(3.72)	(32.19)	(162.65)	(7.46)	(9.92)
Humidity	-90.74	-52.77	-132.56	-252.17	-15.18	5.82
	(34.04)	(6.29)	(15.80)	(96.06)	(13.46)	(18.17)
Heating degree days	-0.17	-0.33	-8.31	-17.11	-0.84	-0.13
	(0.13)	(0.02)	(0.52)	(8.79)	(0.12)	(0.20)
Cooling degree days	-0.75	-1.84	-23.02	-45.54	-2.76	-1.28
	(0.27)	(0.06)	(0.78)	(15.55)	(0.18)	(0.41)
Wind speed	-203.78	303.42	4,972.77	6,463.09	636.33	-100.67
	(116.07)	(21.71)	(355.86)	(5,616.67)	(100.63)	(146.3)
Sunshine	101.40	57.14	459.53	725.58	95.41	46.33
	(32.25)	(5.89)	(47.40)	(178.44)	(21.72)	(27.05)
Coast	977.44	820.58	11,179.37	6,583.3	1,691.95	952.21
	(336.65)	(61.86)	(842.74)	(2,581.37)	(168.49)	(262.15)
Violent crime	2.15	1.04	0.18	0.41	0.02	0.02
	(0.40)	(0.06)	(0.01)	(0.10)	(0.00)	(0.01)
Teacher-pupil ratio	44,408.38	17,039.17	34,934.04	22,919.88	1,975.21	8,005.66†
	(9,818.6)	(1,797.02)	(3,684.29)	(4,912.86)	(1,237.25)	(3,662.89)
Visibility	-7.13	-22.53	-692.30	-888.04	-84.67	-0.59
	(14.69)	(2.76)	(84.77)	(1,020.53)	(12.28)	(14.37)
TSP/PM10	-0.75	-13.67	291.77	522.01	12.92	3.92
	(7.88)	(1.44)	(30.78)	(706.95)	(13.97)	(16.18)
NPDES effluent dischargers	-160.25	-182.48	-18.21	-12.62	-3.17	-3.03
	(63.86)	(11.56)	(1.15)	(4.40)	(2.43)	(2.44)
Landfill waste	-0.23	0.23	-0.04	-0.04	-0.01	0.00
	(0.10)	(0.02)	(0.00)	(0.02)	(0.00)	(0.00)
Superfund sites	-221.67	339.18	1,251.71	1,137.57	187.73	10.19
	(91.32)	(17.37)	(64.09)	(575.92)	(17.54)	(22.97)
Treatment, storage & disposal sites	-121.21	5.54	-29.18	-23.38	-4.55	-0.73
	(3.26)	(0.61)	(3.19)	(226.51)	(0.73)	(1.48)
Central city	1,347.97	-1,018.22	2,828.09	1,654.02	507.38	438.30
	(345.01)	(63.51)	(176.40)	(371.76)	(464.81)	(439.93)
New amenities	N	N	N	N	N	Y

Notes: The full implicit amenity prices are the sum of the annual housing and wage differentials as defined in equation (A-6). Standard errors are shown in parentheses. In order to obtain an annual full implicit price, the product of housing coefficients and mean monthly housing expenditure ($665.47) are multiplied by 12 (months per year) and the wage coefficients are multiplied by the sample means of workers per household (1.75), working hours per week (39.88), and weeks per year worked (45.11). * BBH denotes the rankings from Blomquist, Berger, and Hoehn (1988) and SL are the rankings in Stover and Leven (1992) who use the same data set as BBH, but a different specification for the hedonic estimation. The BHH and SL full implicit amenity prices are adjusted by CPI inflation and reported in terms of 2000 dollars. † The full implicit price for the teacher-pupil ratio includes the coefficient from interacting the teacher-pupil ratio with the share of students attending private schools. (see main text for more discussion of this specification).

the teacher-pupil ratio, acts as a good proxy for the quality of public schools and thus gets capitalised into housing prices also depends the share of students attending private schools.[24] When we account for this effect by including an interaction effect between the teacher-pupil ratio and the percentage of pupils attending private schools in column (6), the full implicit price of school quality increases fourfold compared to model 4 (from $1,975 to $8,005), but is still less than one fifth of the magnitude compared to the estimate of $44,408 in BBH, but half as big as the estimate reported by SL.

Table 3.6: Quality-of-life indices and hedonic specification

		BBH	SL	*BBH amenities* Box-Cox LHS	*BBH amenities* Box-Cox full	*BBH amenities* 2-stage estimation OLS	*BBH amenities* 2-stage estimation FGLS	*New amenities* 2-stage estimation OLS	*New amenities* 2-stage estimation FGLS
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BBH amenities	BBH	1.0000	–	–	–	–	–	–	–
	SL	0.0010	1.0000	–	–	–	–	–	–
	Box-Cox (LHS)	0.0992	0.5380	1.0000	–	–	–	–	–
	Box-Cox full	0.1335	0.7240	-0.1198	1.0000	–	–	–	–
	2-stage OLS	0.1230	0.6671	0.9884	-0.1378	1.0000	–	–	–
	2-stage FGLS	0.1268	0.6877	0.9792	-0.1560	0.9969	1.0000	–	–
New amenities	2-stage OLS	0.1231	0.5580	0.6004	-0.2399	0.6267	0.6300	1.0000	–
	2-stage FGLS	0.1238	0.5612	0.5961	-0.2511	0.6222	0.6263	0.9974	1.0000

Notes: BBH denotes the rankings from Blomquist, Berger, and Hoehn (1988) and SL are the rankings in Stover and Leven (1992) who use the same data set as BBH, but a different specification for the hedonic estimation. The rank correlation between any two quality-of-life indices (QOLI_i and QOLI_j) are computed as Spearman rank correlations.

Overall, the improved two-stage estimation combined with the larger amenity set produces more robust and more plausible estimates of the implicit amenity prices which are used to construct the quality-of-life indices as defined in equation (A-8). Before computing the quality-of-life indices for all geographic samples, it is instructive to analyse the impact of the different hedonic specifications discussed above on the quality-of-life indices for the original sample of 253 counties in BBH. The rank correlations among quality-of-life indices from alternative models are reported in table 3.6. As before we use the results reported by BBH and SL as basis of comparison in columns (1) and (2). Perhaps most remarkably, the rankings produced by the quality-of-life indices in BBH and those in SL bear no resemblance at all, with a rank correlation of close to zero despite using the identical underlying dataset. Our new rankings are slightly more correlated compared to the original BBH with rank correlations ranging from 0.099 for the Box-

[24] In other words, in the presence of private schools which act as substitutes to public schools, the perceived public school quality might become less sensitive to the teacher-pupil ratio (Sonstelie, Brunner, and Ardon, 2000).

Cox specification that fully replicates the BBH set-up to 0.133 for the Box-Cox model with full parameter flexibility. While our rankings bear little resemblance to those in BBH, all indices using the 2000 dataset display consistently high correlations with the SL rankings that use the BBH data from 1980. With a rank correlation of 0.724, our indices produced from the fully-flexible Box-Cox with the restricted amenity set exhibit the most similarity to the SL rankings. Nonetheless, our benchmark two-stage model which uses the full range of amenities still has a rank correlation of 0.561 with the SL rankings, as opposed to a mere 0.124 with the BBH indices. Although underscoring the relative importance for functional specification, these results also identify both temporal changes in the levels of the original amenity variables and the overall bundle of amenities as important sources of variation in the quality of life between counties.

3.7 New quality-of-life rankings

Different amenity endowments across counties create different values of quality-of-life indices for these counties. Because these indices are measured in dollars, they provide a direct measure for the average household's willingness to pay for quality of life in area j as opposed to area k. More specifically, the difference in the value of the index between county j and county k measures the premium that the average household pays implicitly through the housing and labour markets to live in the more amenable county.

3.7.1 Interurban differences in the quality of life

Table 3.7 reports the new rankings by quality of life for the sample of 253 urban counties originally covered by BBH. For each county, we list which metropolitan statistical area (MSA) it belongs to, different new and old rankings and the quality-of-life premium. Column (1) reports the new rank based on the two-stage estimates of the implicit prices for the new set of amenities, also incorporating the improved measure of housing expenditures and after-tax wages. Column (2) reports the rank using the same methodology, but based on the original set of sixteen BHH amenities. Column (3) lists the ranking in BBH, whereas columns (4) and (5) give the z-score and dollar value of the quality-of-life premium that corresponds to the rank in column (1). The last column expresses the implicit quality-of-life expenditure as share of average household income. All quality-of-life index values are rescaled relative $0 for the lowest ranking county in the sample.

The top-ranked county is Marin, California with a $31,513 premium, previously only ranked 142th in the BBH rankings which implied a quality-of-life comparable to that of 143th ranked Comanche County in Oklahoma. Households in second and third placed San Mateo and San Francisco County have the highest implicit quality-of-life expenditures, if measured as share of average household income (35.8%). The index mean for the urban sample is $8,076 with a standard deviation of $5,072 and the county with the (approximate) average quality-of-life index is 100th ranked Medina, Ohio with a $8,066 premium. The least desirable location is Richmond city, which is featured in the top 20 locations in the BBH rankings. Half of the top ten counties were also placed among the ten most amenable places in SL, but none of these locations feature among the top ten places in the BBH rankings. Gaining 223 positions, fourth placed Rockland, New York, a county with one of the highest median household incomes in the nation, records the largest improvement in ranking among the top 25 places.

Table 3.7: Quality-of-life index rankings for BBH urban counties

County	CBSA (MSA or μSA)	Rank			QOLI$_5$ (2000 dollars)*		
		QOLI$_5$	QOLI$_4$	BBH†	z-Score	$ Index	Share‡
		(1)	(2)	(3)	(4)	(5)	(6)
Top 25 counties							
Marin County, CA		1	4	142	4.62	$31,513	0.337
San Mateo County, CA	San Francisco-Oakland-Fremont, CA	2	3	112	4.10	28,896	0.359
San Francisco County, CA		3	1	105	3.76	27,169	0.359
Santa Clara County, CA	San Jose-Sunnyvale-Santa Clara, CA	4	2	88	3.50	25,804	0.304
Nassau County, NY	New York-Northern NJ-Long Island, NY-NJ-PA	5	6	60	2.98	23,169	0.294
Westchester County, NY		6	24	170	2.56	21,048	0.258
Alameda County, CA	San Francisco-Oakland-Fremont, CA	7	5	94	2.50	20,732	0.280
Orange County, CA	Los Angeles-Long Beach-Santa Ana, CA	8	17	41	2.41	20,289	0.293
Bergen County, NJ	New York-Northern NJ-Long Island, NY-NJ-PA	9	10	219	2.24	19,438	0.250
Contra Costa County, CA	San Francisco-Oakland-Fremont, CA	10	11	211	2.23	19,394	0.242
Santa Cruz County, CA	Santa Cruz-Watsonville, CA	11	8	79	2.14	18,934	0.271
Monmouth County, NJ		12	12	92	2.11	18,799	0.237
Rockland County, NY	New York-Northern NJ-Long Island, NY-NJ-PA	13	34	236	2.10	18,727	0.242
New York County, NY		14	7	216	2.06	18,513	0.216
King County, WA	Seattle-Tacoma-Bellevue, WA	15	15	158	1.88	17,588	0.250
Suffolk County, NY	New York-Northern NJ-Long Island, NY-NJ-PA	16	13	49	1.81	17,261	0.236
Ventura County, CA	Oxnard-Thousand Oaks-Ventura, CA	17	22	23	1.71	16,759	0.247
Lake County, IL	Chicago-Naperville-Joliet, IL-IN-WI	18	59	171	1.66	16,498	0.223
Atlantic County, NJ	Atlantic City, NJ	19	23	132	1.58	16,113	0.285
Boulder County, CO	Boulder, CO	20	36	12	1.52	15,785	0.224
Middlesex County, NJ	New York-Northern NJ-Long Island, NY-NJ-PA	21	9	204	1.51	15,734	0.212
Montgomery County, PA	Philadelphia-Camden, PA-NJ-DE-MD	22	30	71	1.44	15,364	0.209
Union County, NJ	New York-Northern NJ-Long Island, NY-NJ-PA	23	68	183	1.42	15,254	0.226
San Diego County, CA	San Diego-Carlsbad-San Marcos, CA	24	38	27	1.36	14,953	0.253
Los Angeles County, CA	Los Angeles-Long Beach-Santa Ana, CA	25	14	58	1.35	14,926	0.253
		⋮			⋮		
Bottom 25 counties							
Nueces County, TX	Corpus Christi, TX	229	170	218	-1.02	2,884	0.061
Wichita County, TX	Wichita Falls, TX	230	69	210	-1.03	2,850	0.064
Cabell County, WV	Huntington-Ashland, WV-KY-OH	231	253	153	-1.11	2,421	0.051

continued on the next page

County	CBSA (MSA or μSA)	Rank QOLI$_5$	QOLI$_4$	BBH[†]	QOLI$_5$ (2000 dollars)[*] z-Score	\$ Index	Share[‡]
		(1)	(2)	(3)	(4)	(5)	(6)
Elkhart County, IN	Elkhart-Goshen, IN	232	149	160	-1.13	2,338	0.044
East Baton Rouge Parish, LA	Baton Rouge, LA	233	240	168	-1.14	2,280	0.041
Saginaw County, MI	Saginaw-Saginaw Township North, MI	234	217	224	-1.14	2,268	0.039
Gaston County, NC	Charlotte-Gastonia-Concord, NC-SC	235	195	76	-1.15	2,256	0.045
Calhoun County, AL	Anniston-Oxford, AL	236	184	159	-1.17	2,162	0.047
Cumberland County, NC	Fayetteville, NC	237	183	95	-1.18	2,086	0.047
Rock County, WI	Janesville, WI	238	130	117	-1.20	1,983	0.035
Norfolk city, VA	Virginia Beach-Norfolk-Newport News, VA-NC	239	29	2	-1.21	1,937	0.046
Hidalgo County, TX	McAllen-Edinburg-Mission, TX	240	252	239	-1.21	1,935	0.056
Allen County, OH	Lima, OH	241	226	228	-1.24	1,801	0.035
Lafayette Parish, LA	Lafayette, LA	242	251	139	-1.27	1,616	0.031
Etowah County, AL	Gadsden, AL	243	192	157	-1.32	1,361	0.029
Yellowstone County, MT	Billings, MT	244	224	187	-1.36	1,195	0.025
Jackson County, MS	Pascagoula, MS	245	222	56	-1.36	1,183	0.024
Jefferson County, TX	Beaumont-Port Arthur, TX	246	179	196	-1.40	973	0.018
Cameron County, TX	Brownsville-Harlingen, TX	247	233	237	-1.40	964	0.027
Scott County, IA	Davenport-Moline-Rock Island, IA-IL	248	132	146	-1.41	928	0.016
Comanche County, OK	Lawton, OK	249	67	143	-1.43	842	0.019
Davidson County, NC	Thomasville-Lexington, NC	250	223	120	-1.45	731	0.015
Spartanburg County, SC	Spartanburg, SC	251	247	57	-1.55	218	0.004
Mobile County, AL	Mobile, AL	252	243	66	-1.56	141	0.003
Richmond city, VA	Richmond, VA	253	153	20	-1.59	0	0

Notes: * The rankings in column (1) correspond to the quality-of-life index values in column (5) which are based on the improved specifications labelled "QOLI$_5$" in the last column of table 3.8. QOLI$_4$ refers to the index based on the original amenities in Blomquist, Berger, and Hoehn (1988) updated for 2000. All QOLI values are rescaled to \$0 for the lowest ranking county in the sample. The mean for the rescaled urban sample is \$8,076 with a standard deviation of \$5,072. This corresponds to an average implicit expenditure share for quality of life of 13% of household income. A complete set of quality of life rankings for all 253 counties is available online at this link. † BBH denotes the original rankings from Blomquist, Berger, and Hoehn (1988). ‡ Implicit quality-of-life expenditure as share of average household income.

Our updated county-level rankings are virtually unrelated to the original BBH rankings. The ranking of eighteen out of our top 25 counties increases by more than 50 places compared to BBH, over a third of those high-rising counties experience a ranking increase of more than 150 positions. Under our preferred specification, only two counties of the highest ranking 25 counties in BBH remain in the top 25: Ventura County, CA (rank 17 vs. rank 23) and Boulder County, CO (rank 20 vs. rank 12). Overall, counties on the West Coast and in larger metro areas are the main beneficiaries of our new quality-of-life index rankings. Among the top 25 places, there are now eleven (previously three) counties from California and one (none in BBH) from Washington State. Furthermore, the top three counties are all in the Bay area: Marin County, San Mateo County and San Francisco County. Counties in the nation's largest metroarea, New York City, are also ranked considerably higher with seven of them now ranked in the top 25. This provides some preliminary evidence against the commonly perceived wisdom of a negative relationship between quality-of-life and urban size (This is addressed more explicitly in section 3.7.3). Nonetheless, these ranking changes are considerably less drastic when compared to the SL rankings – who use the same data as BBH – where sixteen of the highest ranked counties remain in the top 25 places.

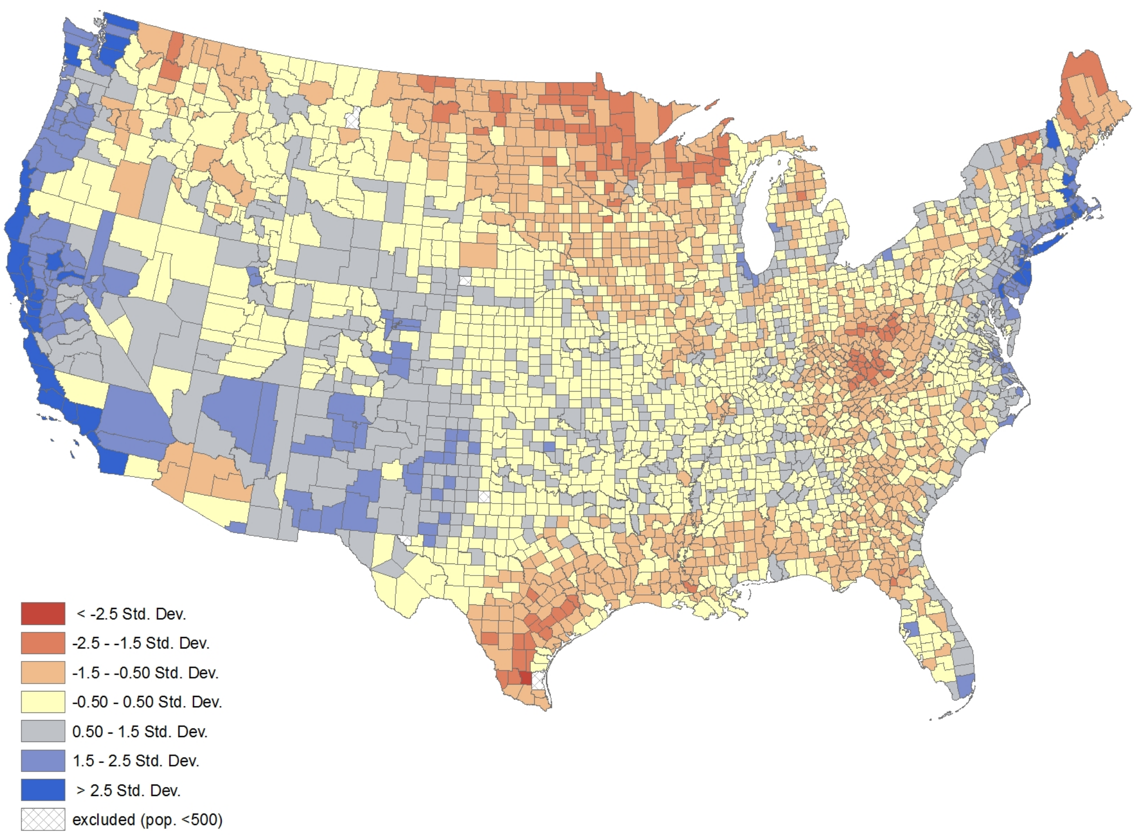
However, the urban counties in BBH cover less than half the total number of metropolitan areas in 2000, only accounting for 58.5% of the approximately 224 million people who live in an urbanised area. In our second geographic sample, we expand coverage to the full set of 1,086 counties that were classified as belonging to one of the 358 MSAs areas in in 2000. Table 3.13 in the appendix reports the rankings and quality-of-life premia for this larger sample of metropolitan counties. In addition to the new rank in column (1), column (2) also lists the matching rank from the smaller sample of urban core counties, where applicable. The top-ranked county is Marin, California as before, except that now the quality-of-life premium relative to the lowest ranked county has grown to $51,824. In this sample, 475th and 476th ranked Fairfield County in South Carolina and Sagadahoc County in Maine represents the locations with a quality of life closest to the average index value of a $34,444 premium. The bottom ranked county is Dawson, Georgia which trails by a quality-of-life discount over $10,500 to the second least desirable county. Furthermore, the bottom 100 counties are separated by a sizeable quality-of-life premium of $30,862 which represents 59.5% of the overall premium between the least and the most amenable metropolitan location. This asymmetry in the quality-of-life variation compared to the smaller sample of BBH counties reflects the larger spread in the amenity endowments across all metropolitan counties. Despite the expanded coverage, however, the quality-of-life ranking remain stable and – with the exception of a few new listings – the rankings in the top 25 remain broadly consistent with those for the BBH sample. There are no new counties in the top ten and Napa and Somoma County in California, and historic Williamsburg in Virginia are among the only six new counties that enter the top 25.

Lastly, we expand our quality-of-life indices to include all counties in the continental US.[25] Table 3.14 reports the rankings for all counties. As before, the top-ranked counties are in the Bay Area with Marin County topping the list with a premium of $72,968. The index mean value for the national sample is $53,294 with a standard deviation of 3,001 and 1634th-ranked Baxter County in Arkansas is the place with the quality-of-life index closest to the sample mean and Fayette County in West Virginia trails the national quality-of-life league table. As in the metropolitan sample, the bottom 100 locations are separated by a significant gap in terms of the quality-of-life premium, this time accounting for an even larger proportion

[25] In order to produce estimates that are consistent with the assumption of a national housing market, the most sparsely populated locations were dropped from the sample. Overall, 29 sparsely-populated rural counties with a population of less than 500 inhabitants are not included in the rankings.

(65.4%) of the overall premium. Given this relatively sizable range of premia – corresponding to over 153% of the national average annual household income – suggests that quality-of-life has become an increasingly important consumption item as measured by its implied budget share. The rankings across our three geographic samples indicate a willingness to pay between 8% to 34% of household income in order to live in the ten most desirable locations in the nation. This compares to only 3% to 9% reported in BBH.

Figure 3.6: Implicit expenditures for quality of life in U.S. counties, 2000



Note: Sparsely populated, rural counties with a population of less than 500 inhabitants are excluded from the sample. *Source*: Authors' calculations.

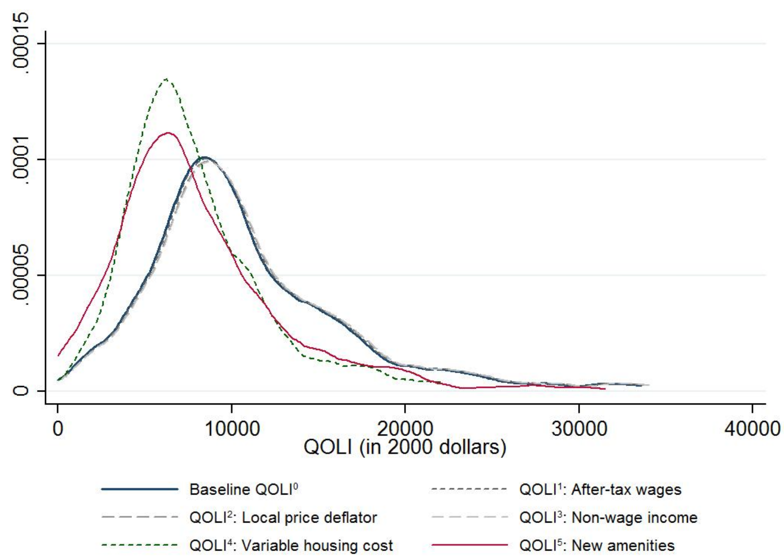
Perhaps the most prominent feature of these rankings is the fact that – with some slight changes in the individual rankings compared with the metropolitan rankings – the top 25 counties are all located in urban areas. Indeed, there are only four counties that are newly ranked in the top 25 and are thus not part of an MSA. Fifth-ranked Lexington city in Virginia is the highest ranked new entry, a small colonial college town which is home to Washington and Lee University, the ninth oldest institution of higher education in the United States, and the Virginia Military Institute, the oldest state-supported military college and one the nation's six senior military colleges. Piktin County in Colorado is another new top 25 entry (rank 8) and – in addition to its scenic beauty and its county seat Aspen being one of the foremost national destinations for wintersports – it is consistently ranked in the top ten highest per-capita income counties in the United States. The full spatial pattern of the national quality-of-life estimates are shown in figure 3.6. The spatial concentrations in the county-level quality-of-life undeniably support the notion that the US truly is a coastal nation (e.g. Rappaport and Sachs, 2003).

3.7.2 Measurement improvements

Aside from differences due to model specification or qualitative variation in the amenity data, quality-of-life rankings are also affected by specific methodological refinements in the context of the housing expenditure and wage data. Recently, the work of Albouy (2008) has re-emphasised wages net of federal taxes and the local variation in the non-housing component of the cost-of-living as important sources for improved estimates in the quality-of-life. Our discussion in section 3.5.3 introduces an additional enhancement by proposing a measure of imputed rents that accounts for the large spatial variation in the user cost of housing. In order to determine their relative importance, we calculate different versions of the quality-of-life indices for each of these improvement for the sample of urban counties and report the impact on the rankings as measured by average change in the rank positions and the corresponding rank correlations.

The baseline specification, QOLI_0, uses amenity data for the year 2000 to replicate the set-up in BBH and is compared to the original BBH rankings and those in SL in columns (1) and (2) of table 3.8 respectively. Compared to the SL rankings, the impact from using updated data on the original amenities changes the ranking of a given county on average by 50 positions, where the largest maximum increase is 175 places and the largest decrease is 155 places respectively. Beginning with models QOLI_1, each stage then successively adds a measurement improvement until the most refined specification, QOLI_5, is reached. The first three specifications, QOLI_1 to QOLI_3, represent the adjustments proposed by Albouy (2008), namely replacing gross income with income net of federal and state taxes, real wages that are

Figure 3.7: Improving QOLI estimates



Notes: Kernel density estimates for different specifications of quality-of-life indices which were computed using the BBH sample and are normalized to be zero in lowest ranked county. QOLI_0 is the base specification and replicates the set-up in BBH with amenity data for the year 2000. QOLI_1 to QOLI_3 represent the adjusted specifications proposed by Albouy (2008), whereas QOLI_4 and QOLI_5 incorporate our improvements.

Table 3.8: Improving quality-of-life estimates

	$QOLI_0^*$ vs. BBH (1)	$QOLI_0^*$ vs. SL (2)	$QOLI_1$ vs. $QOLI_0$ (3)	$QOLI_2$ vs. $QOLI_1$ (4)	$QOLI_3$ vs. $QOLI_2$ (5)	$QOLI_4$ $QOLI_3$ (6)	$QOLI_5$ $QOLI_4$ (7)
Avg. absolute rank change	74	50	1	1	0	10	46
Largest ranking decrease	-223	-155	-2	-4	-2	-57	-182
Largest ranking increase	208	175	4	4	2	44	162
Rank correlation†	0.1471	0.5978	0.9998	0.9998	0.9999	0.9838	0.6286
Federal and state taxes	No		Yes	Yes	Yes	Yes	Yes
Wages	Nominal		Nominal	Real	Real	Real	Real
Non-wage income	No		No	No	Yes	Yes	Yes
User cost of housing	Fixed		Fixed	Fixed	Fixed	Variable	Variable
Amenities	BBH		BBH	BBH	BBH	BBH	New

Notes: BBH denotes the rankings from Blomquist, Berger, and Hoehn (1988) and SL are the rankings in Stover and Leven (1992) who use the same data set as BBH, but a different specification for the hedonic estimation. * $QOLI_0$ is the base specification and replicates the set-up in BBH with amenity data for the year 2000. † Rank correlations use the index in the previous column as reference for comparison. $QOLI_1$ to $QOLI_3$ represent the adjusted specifications proposed by Albouy (2008), whereas $QOLI_4$ and $QOLI_5$ incorporate our improvements.

deflated by a local price index that incorporates non-housing cost and excluding non-wage income).[26] $QOLI_4$ and $QOLI_5$ incorporate the key innovations of paper, i.e. imputed rents that are based on variable user cost and the full set of new amenities.

Compared to the reference specification $QOLI_0$, using the more flexible definition of housing expenditure changes county rankings by an average of ten positions, whereas the data on additional amenities has the largest impact of a ranking change by a further 46 positions on average. With negligible changes on the rank correlations which translates to an average ranking change of no more than one place, using both nominal after-tax wages and real after-tax wages only has a minimal impact on the rankings. The differential contribution of these various improvements to empirical quality-of-life measurements is further illustrated in figure 3.7. The dispersion of the indices produced by separate specification again highlights that the bulk of the changes in the new rankings can be attributed to a more accurate measure of imputed rents and the expanded set of amenities.

While accounting for the local variation in non-housing cost does not appear to have a large impact

[26] To deflate nominal wages by the local cost-of-living, we follow the approach of Moretti (2010) and construct a price index that allows for both the cost of housing and the cost of non-housing consumption to vary across locations area. This approach is discussed in more detail in section 3.A.2. Non-wage income is

on the quality-of-life rankings, it is interesting to note that variations in the local price level (as measured by our local CPI proxy) account for most of the residual variation in our second-stage regressions. Including the CPI with the amenities in the two-stage regression raises the adjusted R^2 to 0.94 in the housing regression and to 0.89 in the wage regression. This is broadly consistent with Winters's (2009) findings that compensating differentials accurately reflect local amenity and cost-of-living differences.

For example, Pueblo County in Colorado – the top ranked county in BBH – falls from position 1 to position 124 in the $QOLI_4$ rankings. If we only used updated data on the sixteen original BBH amenities as in $QOLI_0$, Pueblo County would still be ranked number 20. Using a better measure of user cost of housing pushes its ranking down a further 25 places to number 45, whereas adding new amenities further reduces its ranking to position 124 (down 79 places). The adjustment of wages for federal taxes instead of nominal pre-tax data only has a minimal impact on the rankings. For Pueblo County, after-tax wages improve the ranking by one place to rank 123. However, deflating after-tax wages by the local CPI works in the opposite direction and reduces the after-tax ranking by one place to a final ranking of 124.

Planners, urbanists and economists have a long-standing interest in the relationship between urban size and the resulting quality of life. Traditionally, this discussion is closely linked to the notion of identifying a 'optimal'-sized urban area, i.e. cities where both size-related amenities and disamenities are equally balanced. Although agglomeration effects can be accommodated within the Rosen-Roback framework via productivity effects, the model makes no a priori predictions about the relationship between quality of life and urban size which in the end remains an empirical matter. In contrast to the intuitive appeal of popular rankings, most previously published hedonic quality-of-life indices do not vary with city size or they tend to have a negative relationship.[27] This is documented in table 3.9, listing the most prominent papers that have produced theoretically consistent rankings for which we then compute the relationship between quality-of-life and population size. We also report how these rankings correlate with our own new indices and add the the rankings from the *PlacesRated Almanac* and Sperling's *Cities Ranked & Rated* as a basis for comparison.

In the original BBH rankings, only six counties belonged to a metropolitan area with more than 1.5 million people. Thirteen of the top 25 counties were part of a city with less than half a million inhabitants. In our new rankings, however, small is no longer necessarily beautiful as this pattern is completely reversed: Seventeen counties are part of metropolitan areas with at least three million people and only

[27] Stover and Leven (1992) and Albouy (2008) are the exceptions to this rule.

Table 3.9: Comparison of rank correlations from previous QOLIs

Previous rankings	Index score type	$QOLI_5^*$	$QOLI_0$	City size[†]	Wage weight[‡] areas	Metro
Roback (1982)	Households	0.3873	0.3309	0	–	17
Berger, Blomquist, and Waldner (1987)[♮]	Households	0.2307	0.1448	–	3.61	175
Stover and Leven (1992)[*]	Households	0.4673	0.7112	+	3.61	175
Beeson and Eberts (1988)	Households	-0.2425	-0.3447	0	3.70	34
Gyourko and Tracy (1991)	Households	-0.1323	-0.0506	–	4.82	122
Gabriel and Rosenthal (2004)	Households Firms	0.3519 0.6721	0.3779 0.5488	0 +	–	36
Chen and Rosenthal (2008)	Households Firms	0.0440 0.8205	0.0849 0.5288	– +	2.87	261
Albouy (2008)	Households	0.4637	0.2819	+	1.54	236
PlacesRated Almanac	Crude Adjusted[♭]	0.4734 0.5918	0.5738 0.7173	+ +	–	357
Cities Ranked & Rated	Crude Adjusted[♭]	0.3591 0.6474	0.1133 0.2980	+ +	–	291

Notes: [*] MSA-level QOLIs are obtained from our county-level data by applying the corresponding population weights to the QOLIs of the counties within a given MSA. [†] Within the comparative statics of the Rosen-Roback model, the relationship between city size and quality-of-life depends on the relative interaction between the agglomeration effect and the impact of amenities on individual well-being and on firm productivity. A "+" ("–") indicates that more amenable cities are larger (smaller). A "0" indicates that the relationship is not statistically significant. [‡] Equation A-6 highlights that the full implicit amenity prices (and thus the QOLI) depend on the weight of the wage differential relative to the housing expenditure differential. The wage weight indicates the parameterisation of the relative weight on wages implied by previous studies. Our wage/housing cost weight is 3.19 for urban counties, 3.35 for metropolitan areas and 3.43 for the entire nation. [♮] The rankings in Berger, Blomquist, and Waldner (1987) are the metro-area equivalent of the county-level rankings in Blomquist, Berger, and Hoehn (1988). Their corresponding rank correlation with a population-weighted aggregation of the BBH rankings is 0.9685. [♭] The adjusted scores exclude the job market and cost-of-living components from the crude scores.

three counties belong to small cities where the population is less than half a million.

3.7.3 Intraurban variation in the quality of life

Unlike the bulk of recent study that only consider inter-urban quality-of-life differences at the MSA level, our data permits insights on intra-urban variation in quality-of-life, the main theoretical contribution of BBH. Although urban heterogeneity does not depend on city size alone, it might be difficult to gain meaningful insights on the relative differences in the quality of life between two locations when giving the same observational weight to a one-county metro like Pueblo, Colorado and to a global mega-city like New York City which spans some of the nation's most affluent counties (Hunterdon county, New Jersey and Rockland county, New York) as well as the poorest U.S. Congressional District (Bronx County, NY). Even for large metropolitan areas of similar size such as Atlanta, Boston, or Miami, metro-level aggregates of quality-of-life measures are likely to mask significant intra-urban differences that are influenced by urban structure and population densities: Sprawling over 28 counties, the population density in the Atlanta metroplex is almost three times lower than that in the Boston metropolitan area which covers eight counties, yet almost the same as the density in Miami which covers a mere three counties.

Grouping cities by population size, metropolitan areas with the largest intraurban variation in the quality-of-life at each level of urban scale. The MSA-level aggregate quality-of-life values are calculated as population-weighted averages of the relevant county quality-of-life premia. A set of separate quality-of-life rankings are obtained for the 358 MSAs on the basis of these population-weighted average QOLIs.[28] For the group of the largest US cities with a total population in excess of three million peoples, Atlanta, GA metropolitan area has the largest differential in terms of the quality-of-life premia between the most amenable and least amenable county. Separated by a quality-of-life gap of $38,540, 107th-ranked Fulton County, GA which contains Atlanta is one of the core counties of this large metro, whereas 1,085th-ranked Dawson County, GA is at the outer-fringe of this megaregion that covers 8,376 sq. mi., an area almost the size of New Jersey. The Washington, DC metro area shows similar intra-urban quality-of-life disparities: although less than 50 miles apart, living in 32th-ranked mostly suburban Montgomery county in Maryland as opposed to 1,081st-ranked Jefferson county in West Virginia is worth $19,689 to the average

[28] Interestingly, spatial aggregation matters for the quality-of-life indices. Our rankings have a minimal rank correlation (0.09-0.15) with the original BBH rankings when measured at the county level. However, when we aggregate rankings to the MSA-level, the rank correlation increases to 0.23. This further highlights the importance of intra-urban differences in the quality-of-life which are not captured by studies that use MSA-level aggregates

household.

By contrast, the gradient of intraurban quality-of-life differentiation in the nation's second most desirable MSA, San Francisco which contains first-ranked Marin County, is considerably lower at $6,572. These large differences notwithstanding, the relationship between intraurban heterogeneity does not appear to depend on urban size in a significant manner. Bar a few exceptions, this shows that across all groups of urban scale the most substantial intraurban differences in the quality-of-life are bounded between $2,500 and $10,000. While metropolitan disparities thus seem unrelated to scale, the aggregate MSA quality-of-life indices reveal some positive correlation with urban size, which is broadly consistent with Herzog and Schlottmann (1993) or Albouy (2008). The relationship between urban scale and quality of life is discussed in more detail in chapter 8 which also contains a detailed overview of the largest intraurban quality-of-life differentials in table 8.1.

3.8 National non-market expenditures

Aside from producing mere rankings alone, accurate estimates of implicit household expenditure for non-market goods and services that influence the quality of life also has a broader, macroeconomic motivation. GDP is regularly criticised for its inadequacies of accurately reflecting social welfare and aspects of individual well-being that influence the quality of life.[29] The public debate on alternatives to GDP for the measurement of individual well-being and social welfare has reached a recent high point with the much anticipated recent release of the final report by the "Commission on the Measurement of Economic Performance and Social Progress" (CMEPSP, 2009), called into life by French president, Nicholas Sarkozy, and headed by Joseph Stiglitz and Amartya Sen. The report identifies the reliable measurement and monetary valuation of changes in environmental quality or intangible fixed assets such as public goods as a major hurdle towards a more suitable metric of well-being.

In the United States, for example, the National Income and Product Accounts (NIPA) are sorely lacking when it comes to tracking the value of non-market goods and services. Nordhaus and Tobin (1972) first made this point nearly fourty years ago. Unfortunately, little progress has been made (Nordhaus, 1999, 2000). While NIPA fails to track amenity values, it does track wage income and housing expenditures. This is relevant because there is widespread evidence that spatial variation in rents and wages

[29] Fleurbaey (2009) provides a comprehensive overview of the debate on alternatives to GDP, specifically focussing on different approaches that focus on sustainability, happiness, social choice and fair allocation, and capability.

reflects the values that households assign to environmental services, public goods, and urban activities. Indeed, accurate estimates for personal consumption expenditures on quality of life would complement the corresponding NIPA measures for private goods.

Our estimates seem to confirm the increasing importance of non-market goods and we find that implicit expenditures on non-market goods and services are two to three times larger than previously realized. The average urban household pays between roughly \$8,000 in higher rents and foregone wages to enjoy the amenities that sustain their quality of life. This is equivalent to 13% of gross income as is summarised in the table 3.10 which normalises the quality-of-life indices to be zero in the lowest ranked county. However, our ability to translate quality-of-life differences into implicit real income shares relies heavily on the "no sorting" assumption. Equation (A-6) can also be used to derive a convenient expression of wage and rent differentials and quality-of-life expenditure shares. Re-writing (A-6) as $p_{a_k} da_k = hdr - dw$, dividing by m and re-stating in log-linearised form, we have $\hat{p}_{a_k} = \sigma_h \hat{r} - \sigma_w \hat{w}$, where $\hat{r} \equiv dr/r$, $\hat{r} \equiv dw/w$, $\sigma_h \equiv hr/m$ and $\sigma_w \equiv w/m$. This implies that implied quality of life expenditure is increases with the share of housing expenditure σ_h and decreases with the proportion of wages as part of total household income σ_w.

Table 3.10: Quality-of-life as share of household income

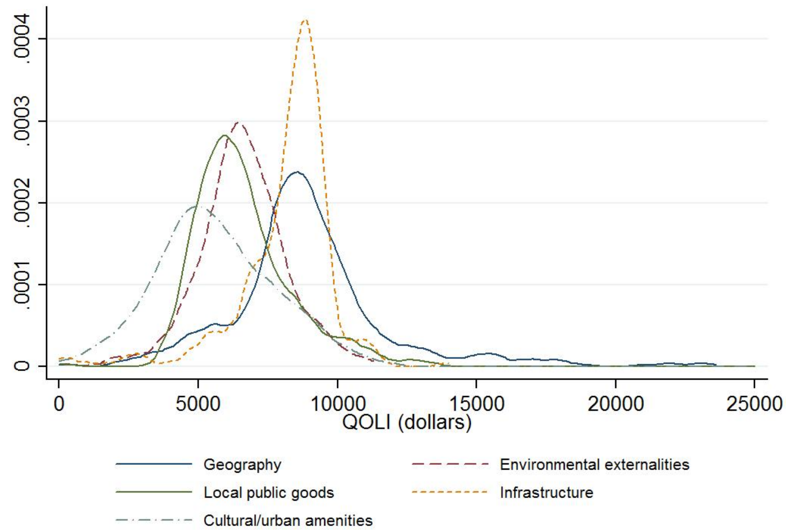
	Sample*	Aggregate QOLI	Income share†	*Climate& geography*	*Environ. externalities*	*Local public goods*	*Infrastructure*	*Cultural & urban amenities*
Mean	BBH	\$8,076	0.133	\$10,358	\$6,520	\$4,137	\$11,539	\$5,959
	Metro	34,250	0.642	5,700	9,220	3,208	8,739	18,621
St. dev.	BBH	5,072	0.065	3,062	1,607	2,266	1,954	2,251
	Metro	3,600	0.055	1,471	1,152	1,449	1,194	2,612
Min	BBH	0	0	0	0	0	0	0
	Metro	0	0	0	0	0	0	0
Max	BBH	31,513	0.359	24,805	11,471	26,770	17,966	12,356
	Metro	51,824	0.974	14,270	13,526	24,958	12,220	44,327

Notes: Quality-of-life index values are normalizes by sub-category to rank counties by each class of amenities. * The BBH sample includes 253 mature urban counties and the metropolitan sample consists of all 1,085 counties that are part of an MSA. † Implicit quality-of-life expenditure as share of average household income.

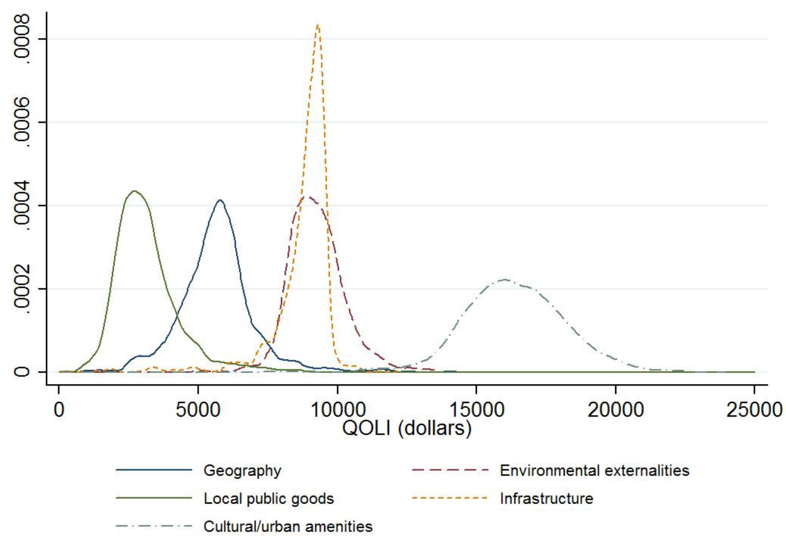
3.8.1 Implicit quality-of-life expenditure by amenity category

Figure 3.8 provides insightful information on relative importance of individual quality-of-life components by illustrating the distribution of implicit household expenditure by amenity category. The most

Figure 3.8: Quality-of-life distributions by amenity category



(a) BBH urban counties



(b) Metropolitan counties

Notes: Kernel density estimates for different specifications of quality-of-life indices. All index values are rescaled to $0 for the lowest ranking county in the sample. The mean for the rescaled urban sample is $8,076 with a standard deviation of $5,072. This corresponds to an average implicit expenditure share for quality of life of 13% of household income. The mean for the rescaled metropolitan sample is $34,249 with a standard deviation of $3,001.

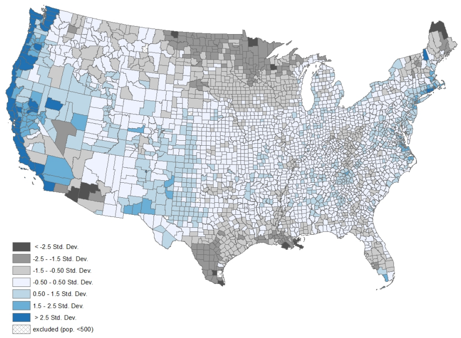
significant category of expenditures in the BBH sample are public infrastructure and climate and geography (e.g. humidity, topography, precipitation) followed by local public goods (e.g. public safety, school quality, municipal parks), environmental externalities (e.g. PM2.5, PM10, NPDES permits, Superfund sites) and cultural and urban activities (e.g. access to restaurants, theatres, golf courses, performing arts). The relative ordering changes for the sample of all metropolitan counties where consumption amenities and the (absence of) environmental externalities attract the highest levels of average implicit household expenditure.

These results provide the foundation for developing a consistent macroeconomic index of amenity value in the United States. They could also be used to adjust cost-of-living measures consistent with Banzhaf's (2005) illustration of the role for weak complementarity in the development of green price indices.

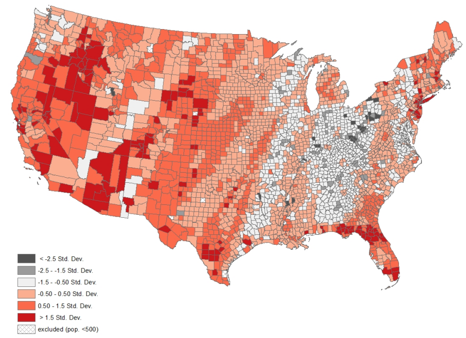
The spatial distribution of the quality-of-life subindices by amenity category for the entire United States is displayed in figure 3.9. Several features of these maps stand out. First, in line with a key tenor in the literature such as Albouy (2008) or Glaeser and Tobio (2008), coastal location, mild seasons and sunshine appear to be the most significant determinant of local differences in quality-of-life expenditures in subpanel 3.9(a). Second, the highest implied expenditure to contain antropogenic externalities coincides with areas that have restrictive environmental regulation and abatement policies, such as the West Coast, or are simply have historically less presence of harmful economic activity. At the same time, the traditionally poorer environmental quality of the Rust Belt is also clearly visible in subpanel 3.9(b). Similarly, subpanel 3.9(c) shows that the largest implied expenditure on local public goods takes places in regions that are well-know for public policy that traditionally focuses on this aspect of quality of life such as the Pacific Northwest, California, and New England. The relative importance of public infrastructure to the most remote areas is strikingly apparent from the map in figure subpanel 3.9(d). At the same time, infrastructure appears to be least valued in terms of its contribution to the quality-of-life in large urban areas. This seems largely consistent with the fact that infrastructure is particularly heavily subsidised in large metropolitan areas where severe congestion externalities regularly outweigh accessability and mobility benefits to residents.

Lastly, there is the category of cultural amenities that includes access activities and services that are traditionally associated with the diversity of larger urban areas such as restaurants, bars, movie theatres, and the performing arts. Indeed, subpanel 3.9(e) shows that such implicit expenditure on these consumption amenities is highest in large metropolitan areas, with particularly high values in Los Angeles

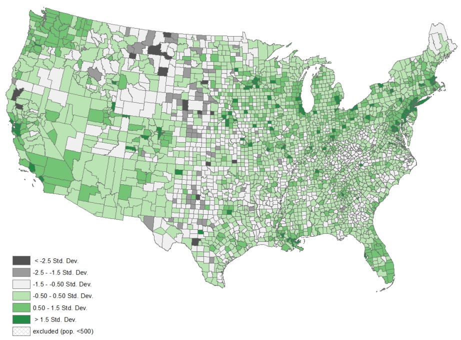
Figure 3.9: Implicit expenditures for quality of life by amenity category, 2000



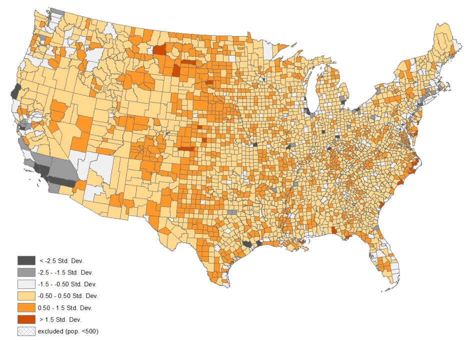
(a) Geography and climate



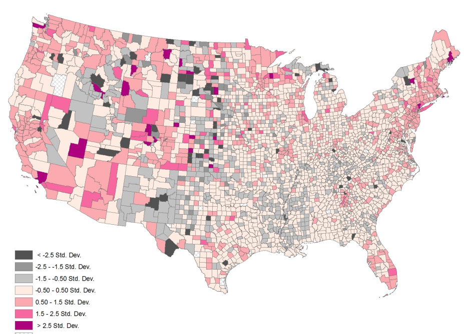
(b) Environmental externalities



(c) Local public goods



(d) Infrastructure



(e) Cultural and urban amenities

Notes: Darker colours indicate higher implicit expenditures by the average household for specific amenity categories. *Geography and climate* includes e.g. land cover diversity, humidity and precipitation; *environmental externalities* includes e.g. PM2.5 and PM10 concentrations, NPDES permits and superfund sites; *local public goods* include e.g. public safety, school quality and municipal parks; *infrastructure* includes urban public transit stops, highway density and railway connectivity; *cultural and urban amenities* include e.g. access to recreational activities, restaurants, bars, movie theatres, and the performing arts. See table 3.2 for a complete list of variables contained in each amenity category.

County and Lincoln County that borders the city of Las Vegas. However, recreational aspects of consumption amenities also account for a significant amount of the quality of life in traditional destinations for tourism and leisure time activities, for example in Pitkin County (Aspen), Routt County (Steamboat Springs) or Summit County (Breckenridge and Copper Mountain) in Colorado, Teton County (Grand Teton National Park and parts of Yellowstone National Park) or Saratoga County (Saratoga Springs and the Adirondacks) in upstate New York.

3.9 Outlook and conclusions

Quality of life matters to households, firms and policy makers alike. Analysing the regional variations in quality-of-life within a framework of compensating differentials provides comprehensive evidence that individuals trade off pecuniary benefits for non-pecuniary improvements in the quality-of-life as measured by higher local amenities in some urban areas.

However, the empirically implemented valuation of amenities using the Rosen-Roback approach is not without problems. More nuanced insights on implied expenditures on non-market goods and services could be obtained by generalising two important aspects of the standard Rosen-Roback model with regard to quality of life. First, there is the need to assess the impact of household heterogeneity on hedonic measures of quality of life. In a first step in this direction, chapter 7 explores how income heterogeneity can be accommodated by adjusting the traditional model. I use a quantile framework to identify evidence of income-related variation in the willingness to pay for amenities. Second, the implications of relaxing the fixed labour supply assumption on quality of life also warrant closer attention. In standard economic theory, labour supply decisions depend on a complete set of prices, not just wages. Local variation in non-wage prices can thus have an impact on households' labour supply decisions. Endogenous labour supply decisions thus introduce an important link between local amenities and differences in local labour markets. The extent to which evidence of local differences in the average weekly hours worked is inconsistent with the assumption of fixing labour supply as a constant in the standard urban location model might yield further insights towards understanding the nature of regional linkages. This is also discussed in more detail in section 7.5 of chapter 7.

Bartik and Smith (1987) caution that public policy needs to be wary of the limiting assumptions that underly the "glorified conversion process view" of assigning a monetary value to consumption amenities which in turn is converted into satisfaction via an individual's residential choice. While research continues to improve these quality of life estimates, policy makers can benefit from these point estimates for valuing amenities. Resource re-allocations from public decisions that ignore these valuations implicitly values amenities anyway, severing the link to economic theory. As part of a theory-informed policy toolkit with obvious limitations, the use of comparative marginal amenity valuations such the quality-of-life indices seems superior to nonuse.

Glaeser and Gottlieb (2009) highlight that modern cities are far more dependent on the role that density can play in speeding the flow of ideas, as opposed to the more traditional view of gains in productivity. In this setting, households in urban areas are likely to have a different valuation of a given bundle of amenities from that of firms. Gabriel and Rosenthal (2004) thus make the useful distinction between "quality of life" and the "quality of the business environment", but they do not look at the relative importance of individual amenities. Determining how different amenities are valued by both firms and households will shed more light on the nature of agglomeration economies across urban areas. Lastly, there is also an interesting case for further endogenising quality-of-life, particularly with regard to closely linked aspects of public policy, such as urban housing policy and land-use regulation, environmental policy, the tax-related benefits to home ownership. Some of these policy-related issues are tentatively explored in chapter 8, but a more comprehensive treatment is left for further research.

3.A Wages and rents

All of our data used in the hedonic wage and housing regressions is taken from the 5% sample of the public use microdata (PUMS) in the 2000 Census. We restrict our sample to non-farm households and person records above the age of 18 for which we construct a measure of hourly wages and monthly housing expenditure.

3.A.1 Hourly wages

Due to the fixed labour supply assumption in the standard Rosen-Roback model, we compute implied hourly wages for full-time workers from self-reported annual income, weeks worked and hours worked per week. Full-time workers are defined using the standard BLS definition as persons who work at least 35 hours or more per week. Using a more restrictive definition of workers who work at least 35 hours a week during at least 26 weeks of the year yields virtually identical results.[30] The summary statistics for the imputed hourly wages across our three samples are reported in table 3.11 below.

In order to assess the impact of regional variations in the burden of federal and state income taxes on quality-of-life estimates, we also derive a measure of hourly after-tax wages. For this purpose, we use estimates of average marginal tax rates for federal and state income taxes for 1999 from the NBER's TAXSIM database (Feenberg and Coutts, 1993). We also account for differences in the level of state excise tax rates which are obtained from from the Book of States (CSG, 2000) minus food tax exemptions (share weighted). The summary statistics of hourly after-tax wages across our three samples are also shown in table 3.11.

3.A.2 Local cost-of-living and non-housing goods

Although the cost of living varies substantially across regions, wage are usually deflated using a single, nation-wide deflator, such as the CPI-U calculated by the BLS. The use a nation-wide deflator is particularly problematic given that more than 40% of the CPI-U is determined by housing costs. The local CPI-U released by the BLS and the ACCRA Cost-of-Living Indices are the two local price indices that are most widely used in empirical work. However, both measures have significant shortcomings: the local CPI-U

[30] The extent to which evidence of local differences in the average weekly hours worked is inconsistent with the assumption of fixing labour supply as a constant is explored in Bieri (2010). He also examines empirical implications for hedonic estimates of quality of life, particularly focusing on previous work that uses microdata while maintaining the fixed labour supply assumption.

is only produced for 23 of the largest metropolitan areas. Furthermore, there are slight differences in the composition of the underlying consumption baskets across cities and the index is normalized to 1 in a given year, thus precluding cross-sectional comparisons. The use of the ACCRA CoLI, on the other hand, might prove problematic due to problems related to its theoretical design, the data collection, and the sampling design, the consequences of which are discussed in more detail in Koo, Phillips, and Sigalla (2000).

The lack of reliable regional cost-of-living indices thus means that most empirical work on quality-of-life does not deflate nominal wages beyond the adjustment in the cost of housing services, as measured by local rents. However, recent work on urban compensating differentials suggests that non-housing goods might also play an important role in determining the local cost-of-living. In order to account for the local variation in the price of non-housing goods, we follow the approach of Moretti (2010) who proposes an index that allows for both the cost of housing and the cost of non-housing consumption to vary across metropolitan areas. While the city-level CPI-U published by the BLS is limited in its geographical coverage, it can still be used to estimate what share of non-housing costs varies with the local cost of housing. The local CPI-U for city j in year t is a weighted average of housing costs (HC_j^t) and non-housing costs (NHC_j^t) such that

$$\mathrm{BLS}_j^t = \alpha \mathrm{HC}_j^t + (1-\alpha)\mathrm{NHC}_j^t, \tag{3.A.1}$$

where α is the CPI weight used by the BLS for housing expenditure. Non-housing costs can now be expressed as consisting of an element that varies systematically with housing costs and an element that evolves independently form housing cost, i.e. $\mathrm{NHC}_j^t = \pi \mathrm{HC}_j^t + \upsilon_j^t$. Using first-differenced prices to avoid non-stationarity then gives the regression set-up $\Delta \mathrm{BLS}_j^t = \beta \Delta \mathrm{HC}_j^t + \varepsilon_j^t$, which in turn can be used to back-out an estimate of π by estimating $\hat{\beta}$, since $\hat{\pi} = \frac{\hat{\beta}-\alpha}{1-\alpha}$. We use panel data on the small sample of 23 MSAs for which a local BLS CPI is available from 1976-2000 to obtain the fixed-effects estimate for β which yields:

$$\Delta \mathrm{BLS}_j = \underset{(0.07)}{1.792} + \underset{(0.01)}{0.619} \; \Delta \mathrm{HC}_j + \varepsilon_j \qquad R^2 = 0.74. \tag{3.A.2}$$

With $\alpha = 0.427$ according to the BLS CPI-U weights in 2000, we can then impute the systematic component of non-housing costs for all MSAs based on their housing cost, i.e. $\hat{\pi}\text{HC}_j^{2000}$ with $\hat{\pi}$ =0.332. Lastly, we compute a local price index as the weighted sum of the cost of housing, the component of non-housing costs that varies with housing, and the component of non-housing costs that does not vary with housing.[31]

3.A.3 Monthly housing expenditure (imputed rents)

We define the dependent variable in the hedonic housing regression as actual rent for renter-occupied units or imputed monthly housing expenditure for owner-occupied units.[32] We drop households with a monthly housing expenditure of less that \$50 from the sample.

Discount rate for imputed rents

Given the pivotal role of the cost of housing (in terms of actual or imputed rents) in determining location-specific quality-of-life estimates, selecting the appropriate discount rate by which housing values are converted into imputed rents is critical. Both Blomquist, Berger, and Hoehn (1988) and Beeson and Eberts (1989) use a fixed discount rate of 7.86% which is based on simulations results by Peiser and Smith (1985, p.357) for an ownership interval from 1987-90 under a scenario of anticipated rising inflation. Despite the secular changes in the interest rate environment that took place over last three decades, more recent studies have paid little attention to this point and use that same constant rate of 7.86% (e.g. Gyourko and Tracy, 1991; Gabriel and Rosenthal, 2004; Albouy, 2008; Chen and Rosenthal, 2008). However, using a uniform discount rate distorts the quality-of-life estimates by ignoring regional variations in housing costs. Ceteris paribus, this bias tends to overrank places which have below average housing cost and underranks locations where housing cost are higher than the national average.

For our imputed rent calculations, we follow the approach of Himmelberg, Mayer, and Sinai (2005) where the annual cost of home ownership C_j^H in location j is defined as

$$C_j^H = P_j r^f + P_j \omega_j + P_j \tau_j (r^m + \omega_j) + P_j \delta_t - P_j \gamma_{t+1} + P_j \varepsilon_t, \tag{3.A.3}$$

[31] Empirically, our parameter estimates are close to Moretti's (2010) estimates of $\hat{\pi}$ = 0.35 which corresponds to $\hat{\beta}$ = 0.63 in 2000. Albouy (2008) uses ACCRA data to run a regression similar to (3.A.2) and obtains a slightly smaller value of $\hat{\pi}$ = 0.26.

[32] It might be argued that observable rental expenditures as opposed (unobservable) imputed rents provide a superior measure of expenditures on housing services. However, Díaz and Luengo-Prado (2008) show that a rental equivalence approach relative to a user cost approach overestimates the cost of housing services.

where

P_j	Value of the property (mid-point of ranges for long-form census variable `VALUE`),
r^f	Risk free rate (10-year average of 3-month T-Bill rate),
r^m	Mortgage rate (10-year average of 30-year fixed rate mortgage),
ω_j	Property tax rate,
τ_j	Marginal income tax rate,
δ_t	Depreciation rate,
γ_{t+1}	Expected capital gain/loss,
ε_t	Owner's risk premium.

Thus the imputed monthly rent can be derived as $R_j = C_j^H/12 = P_j\phi_j$, where the discount factor ϕ_j is defined as

$$\phi_j = 12^{-1} \times (r^f + \omega_j - \tau_j(r^m + \omega_j) + \delta_t - \gamma_{t+1} + \varepsilon_t).^{33} \tag{3.A.4}$$

The discount factor ϕ_t is also known as the user cost of housing and restating $r_j^{\text{imp}} = P_j\phi_j$ as $P_j/r_j^{\text{imp}} = 1/\phi_j$ shows that the equilibrium price-rent ratio should equal the inverse of the user cost. The third term in equation (3.A.4) is particularly worth highlighting, since it represents the tax subsidy to owner-occupied housing due to the tax deductability of mortgage interest and property taxes for households who itemise on their federal taxes.[34]

We compute an implied value for ω_j from the reported property tax payments and house values which has a mean value of 1.64%, 1.59% and 1.54% across the three geographical samples. The summary statistics for ω_j are reported in table 3.12 below. For τ_j, we use average effective marginal income tax rates for 1999 which we collect from the TAXSIM model of the National Bureau of Economic Research. Furthermore, using the empirical estimates from Harding, Rosenthal, and Sirmans (2007), we set r^f = 4.50%, r^m = 5.50%, δ_t = 2.50%, γ_{t+1} = 3.80% (long-run inflation 2.00% plus real appreciation 1.85%) and ε_t = 2.00%.

[33] Equation (3.A.4) is an approximation to the formula for imputed rent in Poterba (1984) who provdes an equilibrium model that links tax deductability of mortgage payments, inflation and home ownership rates.

[34] Since Himmelberg, Mayer, and Sinai (2005) report that less than half of tax-filing homeowners actually itemise and thus do not benefit from these deductions, we reduce the tax subsidy in our calculations by one half. But even without itemising, all homeowners receive some tax subsidy as imputed rents do not have to be reported as taxable income.

Given that the bulk of the spatial variation of user cost comes from at least one of the three components of the tax subsidy to owner-occupied housing (mortgage interest deductability, property tax deductability and tax exemption of income from imputed rent), public policy has an uneven geographical impact on the determinants of quality-of-life. Gyourko and Sinai (2003) calculate the value of the tax subsidy across the nation and report substantial differences across locations in the level of the subsidy flow. At the state level, they report value of total tax code-related benefits per owner-occupied housing unit that range from $917 in South Dakota to $8,092 in California and $8,728 in the District of Columbia. Measured at the metropolitan level, the dispersion is even greater with a low value of $745 (McAllen-Edinberg, TX) and a high of $8,810 (Los Angeles-Riverside-Orange County). Taking in to consideration the impacts of other spatially nonneutral policies, Gyourko and Sinai report that close to 90% of the $26 billion in net transfers between regions goes to the top five metropolitan areas of Los Angeles-Riverside-Orange County, New York-Northern New Jersey, San Francisco-Oakland-San Jose, Boston-Worcester-Lawrence and Washington D.C-Baltimore. These large spatial disparities in tax subsidies to owner-occupied housing are fully consistent with the geographic variation in the user cost of housing represented in figure 3.5.

Dealing with grouped, self-reported housing values

In the long form of the 2000 Census (question 51), housing values are self-reported in 24 intervals from "less than $10,000" to a top-coded category of "$1,000,000 or more". This implies that the data on housing values, our dependent variable for the housing hedonic regressions, is both interval censored and left- and right-censored. Using an ad-hoc OLS regression on the midpoints of the intervals of such grouped data could lead to inconsistent estimates, because it might not adequately reflect the true uncertainty concerning the nature of the exact values within each interval and because it might also inadequately deal with the left- and right-censoring issues in the tails. We address this issue by comparing the parameters from estimating the housing regression via OLS using the interval mid-points to those from using the more appropriate maximum-likelihood interval estimator.

However, largely as a result of our large sample size combined with a relatively high number of intervals (and disturbances which are most likely normally distributed), we do not find a significant differences between the two sets of parameter estimates. This means that the consequences of grouping do not appear to be too severe in our particular case. Furthermore, the root mean-square errors for the two estimators are very similar which suggests that the loss of precision due to using interval midpoints is relatively small and confirms the large-sample findings of Stewart (1983).[35]

Although owners tend to overstate the value of their homes compared to actual sales values, Kiel and Zabel (1997, 1999) provide evidence that the magnitude of the overvaluation is relatively small (≃5%), and – more importantly – that the valuation errors are not systematically related to characteristics of the homeowners, structural characteristics of the house, or the neighbourhood. This implies that empirical estimates based on self-reported house values will provide unbiased estimates of the hedonic prices of both house and amenity characteristics. The summary statistics for the imputed hourly wages across our three samples are reported in table 3.11 below.

3.B Amenity data

Our amenity data comes from a variety of sources which are listed in the last column of table 3.2. All of our data on amenities is collected at the county level and we make two types of transformations or modifications to that data in our analysis.

3.B.1 County-level amenity data

For some of the geographic and environmental amenity data, we use irregularly-spaced spatial NOAA and EPA source data from which we then produce county-level data. For this type of data, we spatially interpolate the amenity data to the population-weighted county centroids via universal kriging. The spatial prediction using universal kriging produces superior results compared to more simple spatial techniques such as inverse distance weighting, because it permits the variogram (i.e. the spatial dependence of the data) to assume a different functional forms that include directional dependence.

3.B.2 PUMAs and counties

Because public use microdata areas (PUMA) that receive the 5-percent data must have a minimum census population of 100,000, the geographic area of PUMAs varies inversely with population density, averaging around 259 people per square mile at the county-level. In most urban areas, there are several PUMAs that fall within a given county, whereas a PUMA can span several counties in the most sparsely

[35] We adjust the top-coded housing values by multiplying them by 1.5.

populated areas. For example, the most densely populated county (New York County, NY) has a population density of 66,951 people per sq.mi and is covered by ten PUMAs. By contrast, in Loving County, TX, which is both the least populous and the least densely populated county in the US, there are only 0.09 people per sq.mi. and its corresponding PUMA covers fourteen counties. The geographical variation in the spatial PUMA-to-county relationship is illustrated in figure 3.10, highlighting that the number of PUMAs per county increases with populations density in urban areas. By design, this relationship is a function of both absolute population size and relative population density and is thus particularly pronounced in the more mature urban areas on the East and West Coast (cf. subfigures 3.10(b) and (c) below).

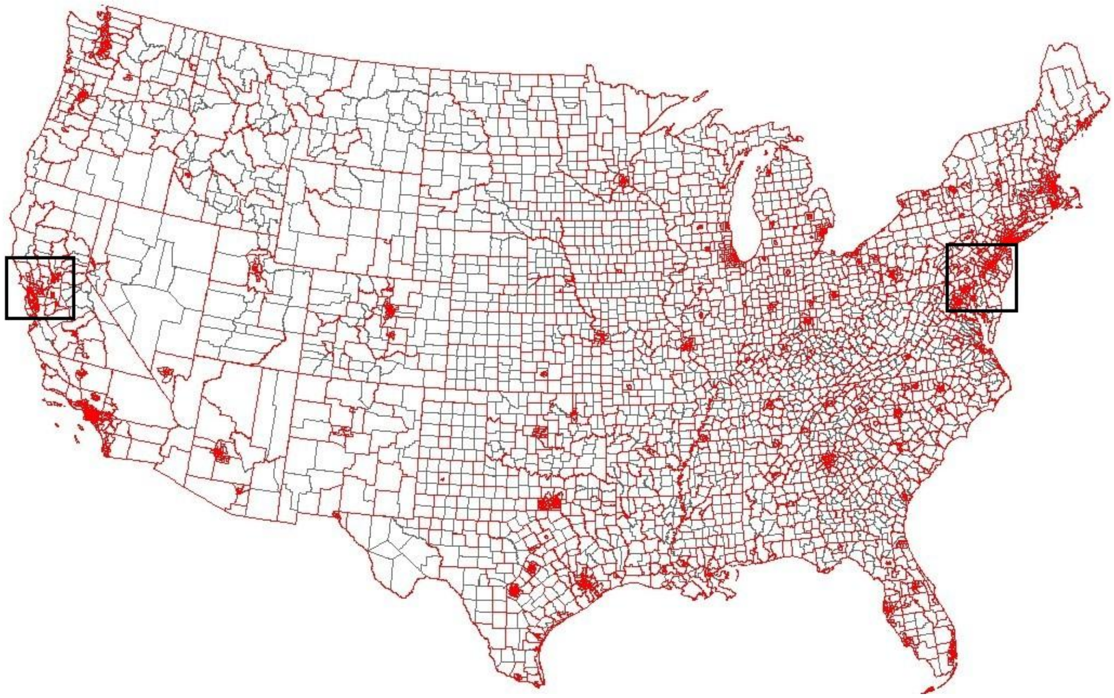
For example, the most populous county is the county with the largest number of PUMAs per county; with a population density of 2,344 people per sq.mi., the 9.51 million people living in Los Angeles County, CA are assigned to 67 PUMAs. The surface area of a PUMA, however, is determined by population density; the average PUMA in New York County covers only 2.2 sq.mi., as opposed to 60.9 sq.mi. in Los Angeles County. Since there is no one-to-one correspondence between PUMAs and counties and virtually all amenity data is only available at the county-level, we assign each PUMA its amenity values by applying the appropriate population weights to the corresponding county-level amenity data. In most cases, there are are multiple PUMAs in each county – approximately four PUMAs per county on average in the BBH sample as opposed to two PUMAs per county in the metropolitan sample (see table 3.3) – and we simply apply unitary weights. For less densely populated areas PUMAs can span several counties, however, and amenities are population-weighted. More formally, this implies that amenity a_k in PUMA j is calculated as

$$a_{kj}^{\text{PUMA}} = \sum_{m=1}^{M} \psi_m a_{kj}^{\text{County}}, \tag{3.B.5}$$

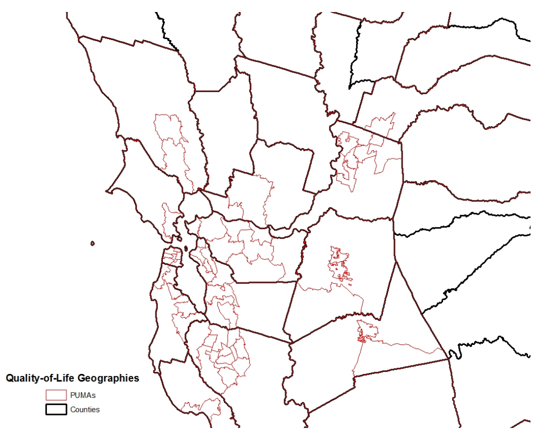
where the weights $\psi_m = p_m / \sum_{m=1}^{M} p_m$ are derived from the 2000 Census population p_m in each of the M counties that exhaustively make up the PUMA. As mentioned above, however, $M = 1$ for virtually all PUMAs in the BBH sample and the metropolitan sample and thus $\psi = 1$.[36]

[36] In the case where $M > 1$, PUMAs only include entire counties and do not reach across state boundaries.

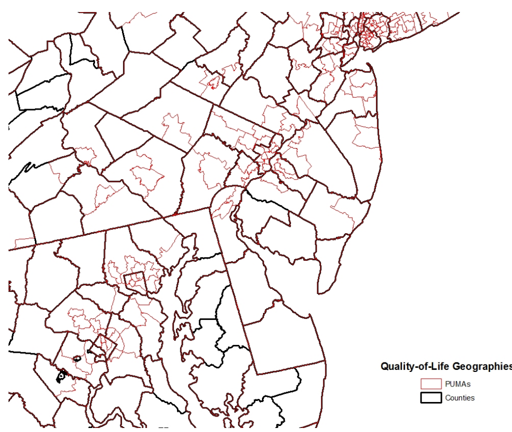
Figure 3.10: Quality-of-life geographies



(a) Counties and PUMAs across the conterminous United States



(b) Counties and PUMAs for the San Francisco Bay Area



(c) Counties and PUMAs for the Capital Region and NYC

Source: Authors' calculations, U.S. Census Bureau.

3.C Supplemental online appendices

Electronic versions of all tables listed in this section, additional figures and a complete set of all quality of life rankings is available online at http://filebox.vt.edu/users/dsbieri/qoli/QOLI_OnlineAppendix.html

3.C.1 Summary statistics

Table 3.11: Person record summary statistics

	Mean	*Std. Dev.*	*Min.*	*Max.*
BBH COUNTIES				
Age	39.48	13.2	18	93
Weeks worked in 1999	45.11	12.7	1	52
Hours per week in 1999	39.93	11.97	5	99
Wage/salary income in 1999	34,591.73	40,794.10	10	347,000
Gross hourly wage	19.02	24.19	1.50	500
Hourly wage (after federal taxes)	14.15	17.98	1.09	385.70
Average marginal federal tax rate (%)	25.59	1.61	20.29	27.51
N. Obs		3,223,602		
MSAs				
Age	39.74	13.35	18	93
Weeks worked in 1999	45.00	12.81	1	52
Hours per week in 1999	39.82	11.95	5	99
Wage/salary income in 1999	32,774.68	38,538.47	10	385,000
Gross hourly wage	18.10	23.05	1.50	500
Hourly wage (after federal taxes)	13.49	17.15	1.09	390.70
Average marginal federal tax rate (%)	25.46	1.63	20.29	27.51
N. Obs		5,827,743		
CONTERMINOUS US				
Age	39.80	13.37	18	93
Weeks worked in 1999	44.89	12.89	1	52
Hours per week in 1999	39.83	12.02	5	99
Wage/salary income in 1999	32,046.98	38,249.52	20	385,000
Gross hourly wage	17.62	22.51	1.50	500
Hourly wage (after federal taxes)	13.17	16.84	1.09	395.95
Average marginal federal tax rate (%)	25.39	1.59	20.29	27.51
N. Obs		6,630,030		

Table 3.12: Housing summary statistics

	Mean	*Std. Dev.*	*Min.*	*Max.*
BBH COUNTIES				
Number of rooms	5.41	2.03	1	9
Number of bedrooms	2.57	1.12	0	5
Acreage	0.86	2.02	0.1	15
Property value	106,632	153,198.1	5,000	1,000,000+
Gross rent	222.59	393.54	4	2,833
Effective property tax rate (%)	1.37	0.94	0	11.49
User cost of housing (%)	4.53	0.65	3.22	13.20
Price-rent ratio	22.08	3.17	31.06	7.58
Monthly housing expenditures ($)	665.47	479.67	50	4,290.42
Workers per household	1.75	1.39	0	4
N. Obs		2,395,116		
MSAS				
Number of rooms	6.18	1.69	1	9
Number of bedrooms	2.98	0.9	0	5
Acreage	1.31	2.80	0.1	15
Property Value	96,201	136,991	5,000	1,000,000+
Gross rent	190.69	358.33	0	2,833
Effective property tax rate (%)	1.28	0.93	0	11.49
User cost of housing (%)	4.47	0.62	3.22	13.20
Price-rent ratio	22.37	3.25	31.06	7.58
Monthly housing expenditures ($)	600.15	463.32	50	3,926.11
Workers per household	1.77	1.38	0	4
N. Obs		4,392,406		
CONTERMINOUS US				
Number of rooms	6.15	1.68	1	9
Number of bedrooms	2.97	0.89	0	5
Acreage	1.52	3.08	0.1	15
Property value	92,535.94	132,544	5,000	1,000,000+
Gross Rent	175.19	340.25	0	2,917
Effective property tax rate (%)	1.28	0.95	0	12.49
User cost of housing (%)	4.48	0.68	3.22	13.20
Price-rent ratio	22.32	3.24	31.06	7.58
Monthly housing expenditures ($)	571.19	450.82	50	3,926.11
Workers per household	1.76	1.38	0	4
N. Obs		5,163,123		

3.D Rankings

3.D.1 Metropolitan counties

Table 3.13: Quality-of-life index rankings for metropolitan counties

County	MSA	Rank MSA (1)	Rank Urban (2)	QOLI (2000 dollars)* z-score (3)	QOLI (2000 dollars)* $ index (4)
Top 25 counties					
Marin County, CA	San Francisco-Oakland-Fremont, CA	1	1	4.88	$51,824
San Francisco County, CA	San Francisco-Oakland-Fremont, CA	2	3	4.56	50,671
San Mateo County, CA	San Francisco-Oakland-Fremont, CA	3	2	4.41	50,127
Santa Clara County, CA	San Jose-Sunnyvale-Santa Clara, CA	4	4	4.11	49,032
Nassau County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	5	5	3.63	47,309
Westchester County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	6	6	3.49	46,805
Contra Costa County, CA	San Francisco-Oakland-Fremont, CA	7	10	3.06	45,260
Alameda County, CA	San Francisco-Oakland-Fremont, CA	8	7	3.06	45,252
Bergen County, NJ	New York-Northern New Jersey-Long Island, NY-NJ-PA	9	9	3.06	45,249
Rockland County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	10	13	3.03	45,164
Carson City, NV	Carson City, NV	11	–	3.02	45,114
Napa County, CA	Napa, CA	12	–	3.00	45,059
Suffolk County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	13	16	2.92	44,768
Orange County, CA	Los Angeles-Long Beach-Santa Ana, CA	14	8	2.92	44,755
Middlesex County, MA	Boston-Cambridge-Quincy, MA-NH	15	–	2.84	44,490
Monmouth County, NJ	New York-Northern New Jersey-Long Island, NY-NJ-PA	16	12	2.75	44,147
Santa Cruz County, CA	Santa Cruz-Watsonville, CA	17	11	2.58	43,538
New York County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	18	14	2.57	43,513
Morris County, NJ	New York-Northern New Jersey-Long Island, NY-NJ-PA	19	–	2.57	43,505
Sonoma County, CA	Santa Rosa-Petaluma, CA	20	–	2.46	43,115
King County, WA	Seattle-Tacoma-Bellevue, WA	21	15	2.44	43,024
Middlesex County, NJ	New York-Northern New Jersey-Long Island, NY-NJ-PA	22	21	2.36	42,758
Williamsburg city, VA	Virginia Beach-Norfolk-Newport News, VA-NC	23	–	2.33	42,643
Ventura County, CA	Oxnard-Thousand Oaks-Ventura, CA	24	17	2.31	42,581
Lake County, IL	Chicago-Naperville-Joliet, IL-IN-WI	25	18	2.30	42,527
		⋮		⋮	
Median 25 counties					
Winnebago County, WI	Oshkosh-Neenah, WI	531	–	-0.08	33,957
Warrick County, IN	Evansville, IN-KY	532	–	-0.08	33,955
Pender County, NC	Wilmington, NC	533	–	-0.09	33,928
Forsyth County, NC	Winston-Salem, NC	534	186	-0.09	33,917
Hanover County, VA	Richmond, VA	535	–	-0.09	33,916
Davie County, NC	Winston-Salem, NC	536	–	-0.09	33,916
Asotin County, WA	Lewiston, ID-WA	537	–	-0.09	33,914
Franklin County, MO	St. Louis, MO-IL	538	–	-0.10	33,907
Grayson County, TX	Sherman-Denison, TX	539	–	-0.10	33,893
Shawnee County, KS	Topeka, KS	540	–	-0.10	33,889
Erie County, OH	Sandusky, OH	541	–	-0.10	33,881
Pitt County, NC	Greenville, NC	542	–	-0.10	33,873
Whitley County, IN	Fort Wayne, IN	543	–	-0.10	33,872
Clark County, IN	Louisville-Jefferson County, KY-IN	544	–	-0.11	33,855

continued on the next page

County	MSA	Rank		QOLI (2000 dollars)*	
		MSA	Urban	*z*-score	$ index
		(1)	(2)	(3)	(4)
Hampshire County, WV	Winchester, VA-WV	545	–	-0.11	33,853
Campbell County, VA	Lynchburg, VA	546	–	-0.11	33,841
Shelby County, TN	Memphis, TN-MS-AR	547	174	-0.11	33,838
Somerset County, MD	Salisbury, MD	548	–	-0.12	33,824
Charles County, MD	Washington-Arlington-Alexandria, DC-VA-MD-WV	549	–	-0.12	33,824
Franklin County, ID	Logan, UT-ID	550	–	-0.12	33,821
Madison County, AL	Huntsville, AL	551	–	-0.12	33,819
Williamson County, TX	Austin-Round Rock, TX	552	–	-0.13	33,795
Richmond County, GA	Augusta-Richmond County, GA-SC	553	215	-0.13	33,793
Washington County, MO	St. Louis, MO-IL	554	–	-0.13	33,785
Bartholomew County, IN	Columbus, IN	555	–	-0.13	33,780
Bottom 25 counties		⋮		⋮	
Guthrie County, IA	Des Moines-West Des Moines, IA	1,061	–	-1.53	28,737
Jefferson County, FL	Tallahassee, FL	1,062	–	-1.53	28,728
Richmond city, VA	Richmond, VA	1,063	253	-1.54	28,700
George County, MS	Pascagoula, MS	1,064	–	-1.55	28,676
Power County, ID	Pocatello, ID	1,065	–	-1.58	28,561
Marshall County, IL	Peoria, IL	1,066	–	-1.67	28,237
Currituck County, NC	Virginia Beach-Norfolk-Newport News, VA-NC	1,067	–	-1.70	28,118
Logan County, OK	Oklahoma City, OK	1,068	–	-1.71	28,082
Boyd County, KY	Huntington-Ashland, WV-KY-OH	1,069	–	-1.77	27,861
Pleasants County, WV	Parkersburg-Marietta-Vienna, WV-OH	1,070	–	-2.06	26,848
Newton County, GA	Atlanta-Sandy Springs-Marietta, GA	1,071	–	-2.07	26,802
Colonial Heights city, VA	Richmond, VA	1,072	–	-2.23	26,218
Morton County, ND	Bismarck, ND	1,073	–	-2.24	26,173
Calhoun County, IL	St. Louis, MO-IL	1,074	–	-2.36	25,749
Washington County, OH	Parkersburg-Marietta-Vienna, WV-OH	1,075	–	-2.54	25,116
Lawrence County, OH	Huntington-Ashland, WV-KY-OH	1,076	–	-2.60	24,881
Walker County, GA	Chattanooga, TN-GA	1,077	–	-2.79	24,214
Greenup County, KY	Huntington-Ashland, WV-KY-OH	1,078	–	-2.80	24,181
Bristol city, VA	Kingsport-Bristol-Bristol, TN-VA	1,079	–	-2.80	24,161
Wood County, WV	Parkersburg-Marietta-Vienna, WV-OH	1,080	–	-2.89	23,859
Jefferson County, WV	Washington-Arlington-Alexandria, DC-VA-MD-WV	1,081	–	-3.43	21,898
Storey County, NV	Reno-Sparks, NV	1,082	–	-4.52	17,990
Goliad County, TX	Victoria, TX	1,083	–	-4.94	16,481
Polk County, TN	Cleveland, TN	1,084	–	-6.58	10,577
Dawson County, GA	Atlanta-Sandy Springs-Marietta, GA	1,085	–	-9.51	0

Notes: * The rankings in column (1) correspond to the quality-of-life index values in column (4) which are based on the improved specifications labelled "$QOLI_5$" in the last column of table 3.8. All QOLI values are rescaled to $0 for the lowest ranking county in the sample. The mean for the rescaled metropolitan sample is $34,294 with a standard deviation of $3,001. A complete set of quality of life rankings for all 1,085 metropolitan counties is available online at this link.

3.D.2 All counties

Table 3.14: Quality-of-life index rankings for all counties

County	CBSA (MSA or μSA)	Rank All	Rank MSA	QOLI (2000 dollars)* z-score	QOLI (2000 dollars)* $ index
		(1)	(2)	(3)	(4)
Top 25 counties					
Marin County, CA		1	1	1.88	$72,968
San Francisco County, CA	San Francisco-Oakland-Fremont, CA	2	2	1.84	72,550
San Mateo County, CA		3	3	1.73	71,365
Santa Clara County, CA	San Jose-Sunnyvale-Santa Clara, CA	4	4	1.70	71,054
Lexington city, VA		5	–	1.55	69,537
Nassau County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	6	5	1.41	68,026
Westchester County, NY		7	6	1.40	67,847
Pitkin County, CO		8	–	1.34	67,220
Alameda County, CA	San Francisco-Oakland-Fremont, CA	9	8	1.30	66,873
Contra Costa County, CA		10	7	1.30	66,803
Orange County, CA	Los Angeles-Long Beach-Santa Ana, CA	11	14	1.29	66,700
Jefferson County, WA		12	–	1.24	66,179
Napa County, CA	Napa, CA	13	12	1.24	66,175
Rockland County, NY		14	10	1.24	66,153
Suffolk County, NY	New York-Northern New Jersey-Long Island, NY-NJ-PA	15	13	1.22	65,994
New York County, NY		16	18	1.21	65,904
Bergen County, NJ		17	9	1.21	65,881
Middlesex County, MA	Boston-Cambridge-Quincy, MA-NH	18	15	1.18	65,545
Monmouth County, NJ	New York-Northern New Jersey-Long Island, NY-NJ-PA	19	16	1.16	65,329
Tillamook County, OR		20	–	1.11	64,802
Sonoma County, CA	Santa Rosa-Petaluma, CA	21	20	1.09	64,584
Middlesex County, NJ	New York-Northern New Jersey-Long Island, NY-NJ-PA	22	22	1.08	64,444
Williamsburg city, VA	Virginia Beach-Norfolk-Newport News, VA-NC	23	23	1.07	64,377
Santa Cruz County, CA	Santa Cruz-Watsonville, CA	24	17	1.05	64,187
Carson City, NV	Carson City, NV	25	11	1.05	64,183
		⋮		⋮	
Median 25 counties					
Shannon County, MO		1,528	–	0.02	53,204
Taylor County, WI		1,529	–	0.02	53,204
Roseau County, MN		1,530	–	0.02	53,203
Marion County, OH	Marion, OH	1,531	–	0.02	53,200
Clark County, MO	Fort Madison-Keokuk, IA-MO	1,532	–	0.02	53,197
Lenawee County, MI	Adrian, MI	1,533	–	0.02	53,197
Crawford County, KS	Pittsburg, KS	1,534	–	0.02	53,189
Pittsylvania County, VA	Danville, VA	1,535	673	0.02	53,188
Hancock County, KY	Owensboro, KY	1,536	598	0.02	53,188
Screven County, GA		1,537	–	0.02	53,187
Hardee County, FL	Wauchula, FL	1,538	–	0.02	53,186
Columbia County, GA	Augusta-Richmond County, GA-SC	1,539	710	0.02	53,186
Richland County, ND	Wahpeton, ND-MN	1,540	–	0.02	53,185
Onslow County, NC	Jacksonville, NC	1,541	801	0.02	53,183
Jay County, IN		1,542	–	0.02	53,183
McCurtain County, OK		1,543	–	0.02	53,180
Kewaunee County, WI	Green Bay, WI	1,544	924	0.02	53,179
Pickaway County, OH	Columbus, OH	1,545	638	0.02	53,179

continued on the next page

County	CBSA (MSA or μSA)	Rank		QOLI (2000 dollars)*	
		All	MSA	z-score	$ index
		(1)	(2)	(3)	(4)
Waller County, TX	Houston-Sugar Land-Baytown, TX	1,546	695	0.02	53,178
Simpson County, MS	Jackson, MS	1,547	558	0.02	53,178
Tippecanoe County, IN	Lafayette, IN	1,548	802	0.01	53,177
Harnett County, NC	Dunn, NC	1,549	–	0.01	53,172
Fairfield County, OH	Columbus, OH	1,550	728	0.01	53,172
Northumberland County, PA	Sunbury, PA	1,551	–	0.01	53,171
Hart County, GA		1,552	–	0.01	53,168
		⋮		⋮	
Bottom 25 counties					
Perkins County, SD		3,048	–	-1.24	39,871
Daniels County, MT		3,049	–	-1.29	39,312
Trego County, KS		3,050	–	-1.38	38,366
Polk County, TN	Cleveland, TN	3,051	1,084	-1.38	38,346
Hyde County, SD		3,052	–	-1.45	37,585
Shackelford County, TX		3,053	–	-1.45	37,579
Mackinac County, MI		3,054	–	-1.46	37,528
Emporia city, VA		3,055	–	-1.46	37,525
Jackson County, CO		3,056	–	-1.47	37,389
Sierra County, CA		3,057	–	-1.48	37,345
Boyd County, NE		3,058	–	-1.49	37,203
Lancaster County, SC	Lancaster, SC	3,059	–	-1.50	37,069
Montmorency County, MI		3,060	–	-1.52	36,819
Judith Basin County, MT		3,061	–	-1.72	34,725
Dawson County, GA	Atlanta-Sandy Springs-Marietta, GA	3,062	1,085	-1.95	32,265
Hamilton County, NY		3,063	–	-1.97	32,091
Custer County, ID		3,064	–	-2.27	28,900
Sherman County, OR		3,065	–	-2.28	28,776
Storey County, NV	Reno-Sparks, NV	3,066	1,082	-2.67	24,662
Brewster County, TX		3,067	–	-2.68	24,522
Powder River County, MT		3,068	–	-2.76	23,672
Owen County, KY		3,069	–	-2.83	22,947
Coosa County, AL	Alexander City, AL	3,070	–	-4.39	6,382
Choctaw County, OK		3,071	–	-4.50	5,169
Fayette County, WV	Oak Hill, WV	3,072	–	-4.99	0

Notes: * The rankings in column (1) correspond to the quality-of-life index values in column (4) which are based on the improved specifications labelled "$QOLI_5$" in the last column of table 3.8. All QOLI values are rescaled to $0 for the lowest ranking county in the sample. The mean for the rescaled full sample is $53,395 with a standard deviation of $3,803. 29 rural counties with a population of less than 1,000 inhabitants are not included in the sample. A complete set of quality of life rankings for all 3,072 counties is available online at this link.

3.E Acknowledgements

We would like to thank, without implicating, Kerry Smith and Chris Timmins for useful suggestions and comments from seminar participants at the School of Sustainability at Arizona State University, Camp Resources XVI at North Carolina State University, Portland State University, the Southern Regional Science Association, the University of Michigan and Virginia Tech are also gratefully acknowledged.

3.F References

ALBOUY, D. Y. (2008): "Are Big Cities Really Bad Places to Live? Improving Quality-of-Life Estimates across Cities," Working Paper No. 14472, National Bureau of Economic Research, Cambridge, MA.

——— (2009a): "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 117(4), 635–667.

——— (2009b): "What Are Cities Worth? Land Rents, Local Productivity, and the Capitalization of Amenity Values," Working Paper No. 14981, National Bureau of Economic Research, Cambridge, MA.

ANAS, A., AND D. PINES (2008): "Anti-sprawl Policies in a System of Congested Cities," *Regional Science and Urban Economics*, 38(5), 408–423.

BANZHAF, H. S. (2005): "Green Price Indices," *Journal of Environmental Economics and Management*, 49(2), 262–280.

BANZHAF, H. S., AND N. LAVERY (2010): "Can the Land Tax Help Curb Urban Sprawl? Evidence from Growth Patterns in Pennsylvania," *Journal of Urban Economics*, 67, 169–179.

BARTIK, T. J., AND V. K. SMITH (1987): *Handbook of Regional and Urban Economics: Urban Economics* vol. 2 of *Handbooks in Economics*, chap. Urban Amenities and Public Policy, pp. 1207–1254. North-Holland, Amsterdam.

BAUM-SNOW, N. (2007a): "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 122(2), 775–805.

——— (2007b): "Suburbanization and Transportation in the Monocentric Model," *Journal of Urban Economics*, 62(3), 405–423.

BAYER, P., AND C. TIMMINS (2007): "Estimating Equilibrium Models of Sorting Across Locations," *Economic Journal*, 117(2), 353–374.

BEESON, P. E., AND R. W. EBERTS (1988): "Interurban Comparisons of the Quality of Life," Working Paper No. 8809, Federal Reserve Bank of Cleveland, Cleveland, OH.

——— (1989): "Identifying Productivity and Amenity Effects in Interurban Wage Differentials," *Review of Economics and Statistics*, 71(3), 443–452.

BERGER, M. C., G. C. BLOMQUIST, AND W. WALDNER (1987): "A Revealed-Preference Ranking of Quality-of-Life in Metropolitan Areas," *Social Science Quarterly*, 68(4), 761–778.

BIERI, D. S. (2010): "Location Choice, Linkages and the Spatial Economy: Theory, Evidence and Heterodox Assessment," Ph.D. thesis, School of Public & International Affairs, Virginia Tech, Blacksburg, VA.

BLOMQUIST, G. C. (2006): *A Companion to Urban Economics* chap. Measuring Quality of Life, pp. 483–502, Blackwell Companions to Contemporary Economics. Blackwell Publishing, New York, NY.

BLOMQUIST, G. C., M. C. BERGER, AND J. P. HOEHN (1988): "New Estimates of Quality of Life in Urban Areas," *American Economic Review*, 78(1), 89–107.

BRASINGTON, D. M. (1999): "Which Measures of School Quality Does the Housing Market Value?," *Journal of Real Estate Research*, 18(3), 395–414.

BURCHFIELD, M., H. G. OVERMAN, D. PUGA, AND M. A. TURNER (2006): "Causes of Sprawl: A Portrait from Space," *Quarterly Journal of Economics*, 121(3), 587–632.

CHEN, Y., AND S. S. ROSENTHAL (2008): "Local Amenities and Life-cycle Migration: Do People Move for Jobs or Fun?," *Journal of Urban Economics*, 64(4), 519–537.

CMEPSP (2009): *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Commission on the Measurement of Economic Performance and Social Progress, Paris, France.

COMBES, P.-P., G. DURANTON, AND L. GOBILLON (2008): "Spatial Wage Disparities: Sorting Matters!," *Journal of Urban Economics*, 63(2), 723–747.

CROPPER, M. L., L. B. DECK, AND K. E. MCCONNELL (1988): "On the Choice of Funtional Form for Hedonic Price Functions," *Review of Economics and Statistics*, 70(4), 668–675.

CSG (2000): *The Book of the States*, vol. 33. The Council of State Governments, Lexington, KY.

DEITZ, R., AND J. R. ABEL (2008): "Have Amenities Become Relatively More Important than Firm Productivity Advantages in Metropolitan Areas?," Staff Report No. 344, Federal Reserve Bank of New York, Buffalo, NY.

DÍAZ, A., AND M. J. LUENGO-PRADO (2008): "On the User Cost and Homeownership," *Review of Economic Dynamics*, 11(3), 584–613.

DOWNES, T. A., AND S. M. GREENSTEIN (1996): "Understanding the Supply Decisions of Nonprofits: Modelling the Location of Private Schools," *RAND Journal of Economics*, 27(2), 365–390.

DOWNES, T. A., AND J. E. ZABEL (2002): "The Impact of School Characteristics on House Prices: Chicago 1987–1991," *Journal of Urban Economics*, 52(1), 1–24.

EIU (2005): "The Economist Intelligence Unit's Quality-of-Life Index," The world in 2005, Economist Intelligence Unit, London, UK.

EVANS, M. F., AND V. K. SMITH (2005): "Do New Health Conditions Support MortalityŰAir Pollution Effects," *Journal of Environmental Economics and Management*, 50(3), 496Ű–518.

FEENBERG, D., AND E. COUTTS (1993): "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 12(1), 189–194.

FLEURBAEY, M. (2009): "Beyond GDP: The Quest for a Measure of Social Welfare," *Journal of Economic Literature*, 47(4), 1029–75.

FLORIDA, R. (2008): *Who's Your City? How the Creative Economy Is Making Where to Live the Most Important Decision of Your Life*. Basic Books, New York, NY.

GABRIEL, S. A., J. P. MATTEY, AND W. L. WASCHER (2003): "Compensating Differentials and Evolution in the Quality-of-Life among U.S. States," *Regional Science and Urban Economics*, 33(5), 619–647.

GABRIEL, S. A., AND S. S. ROSENTHAL (2004): "Quality of the Business Environment Versus Quality of Life: Do Firms and Households Like the Same Cities?," *Review of Economics and Statistics*, 86(1), 438–444.

GLAESER, E. L., AND J. D. GOTTLIEB (2006): "Urban Resurgence and the Consumer City," *Urban Studies*, 43(8), 1275–1299.

——— (2009): "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, 47(4), 983Ű–1028.

GLAESER, E. L., J. KOLKO, AND A. SAIZ (2001): "Consumer City," *Journal of Economic Geography*, 1(1), 27–50.

GLAESER, E. L., AND C. TOBIO (2008): "The Rise of the Sunbelt," *Southern Economic Journal*, 74(3), 610Ű–643.

GYOURKO, J. E., M. E. KAHN, AND J. TRACY (1999): *Handbook of Regional and Urban Economics: Applied Urban Economics* vol. 3 of *Handbooks in Economics*, chap. Quality of Life and Environmental Comparisons, pp. 1413–1454. North-Holland, Amsterdam.

GYOURKO, J. E., AND T. SINAI (2003): "The Spatial Distribution of Housing-Related Ordinary Income Tax Benefits," *Real Estate Economics*, 31(4), 527–575.

GYOURKO, J. E., AND J. TRACY (1991): "The Structure of Local Public Finance and the Quality of Life," *Journal of Political Economy*, 99(4), 774–806.

HARDING, J. P., S. S. ROSENTHAL, AND C. F. SIRMANS (2007): "Depreciation of Housing Capital, Maintenance, and House Price Inflation: Estimates from a Repeat Sales Model," *Journal of Urban Economics*, 61(2), 193–217.

HAURIN, D. R. (1980): "The Regional Distribution of Population, Migration, and Climate," *Quarterly Journal of Economics*, 95(2), 293–308.

HERZOG, HENRY W, J., AND A. M. SCHLOTTMANN (1993): "Valuing Amenities and Disamenities of Urban Scale: Can Bigger be Better?," *Journal of Regional Science*, 2(2), 145–165.

HIMMELBERG, C., C. MAYER, AND T. SINAI (2005): "Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions," *Journal of Economic Perspectives*, 19(4), 67–92.

HOEHN, J. P., M. C. BERGER, AND G. C. BLOMQUIST (1987): "A Hedonic Model of Interregional Wages, Rents and Amenity Values," *Journal of Regional Science*, 27(4), 605–620.

HWANG, H.-S., W. R. REED, AND C. HUBBARD (1992): "Compensating Wage Differentials and Unobserved Productivity," *Journal of Political Economy*, 100(4), 835–858.

ICPSR (2008): "U.S. County Characteristics, 2000–2007 [Computer File]," ICPSR Study No. 20660, Inter-university Consortium for Political and Social Research, Ann Arbor, MI.

KAHN, M. E. (1995): "A Revealed Preference Approach to Ranking City Quality of Life," *Journal of Urban Economics*, 38(2), 221–235.

KENNEDY, P. E. (1981): "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 71(4), 801.

KIEL, K. A., AND J. E. ZABEL (1997): "Evaluating the Usefulness of the American Housing Survey for Creating House Price Indices," *Journal of Real Estate Finance and Economics*, 14(1–2), 189–Ű202.

——— (1999): "The Accuracy of Owner-Provided House Values: The 1978Ű1991 American Housing Survey," *Real Estate Economics*, 27(2), 263–298.

KOO, J., K. R. PHILLIPS, AND F. D. SIGALLA (2000): "Mesuring Regional Cost of Living," *Journal of Business & Economic Statistics*, 18(1), 127–136.

KRUPKA, D. J. (2008): "On Amenities, Natural Advantage and Agglomeration," Discussion Paper No. 3598, Institute for the Study of Labor (IZA), Bonn, Germany.

KUMINOFF, N. V., C. F. PARMETER, AND J. C. POPE (2010): "Which Hedonic Models Can We Trust to Recover the Marginal Willingness to Pay for Environmental Amenities?," *Journal of Environmental Economics and Management*, forthcoming.

LAMBIRI, D., B. BIAGI, AND V. ROYUELA (2007): "Quality-of-Life in the Economic and Urban Economic Literature," *Social Indicators Research*, 84(1), 1–25.

LEE, S. C. (2010): "Ability Sorting and Consumer City," *Journal of Urban Economics*, 68(1), 20–33.

MERCER (2009): "2009 Quality of Living Survey," Discussion paper, Marsh and McLennan Companies, London, UK.

MORETTI, E. (2010): "Real Wage Inequality," Mimeograph, University of California at Berkeley, Berkely, CA.

MOULTON, B. R. (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics*, 72(2), 334–338.

NORDHAUS, W. D. (1999): "Beyond the CPI: An Augmented Cost-of-Living Index," *Journal of Business & Economic Statistics*, 17(2), 182–187.

——— (2000): "New Directions in National Economic Accounting," *American Economic Review*, 90(2), 259–263.

NORDHAUS, W. D., AND J. TOBIN (1972): *The Measurement of Economic and Social Performance, Studies in Income and Wealth* vol. 38, chap. Is Growth Obsolete?, pp. 509–564. NBER, Cambridge, MA.

PARTRIDGE, M. D., D. S. RICKMAN, K. ALI, AND M. R. OLFERT (2009): "Agglomeration Spillovers and Wage and Housing Cost Gradients Across the Urban Hierarchy," *Journal of International Economics*, 78(2), 126–140.

PEISER, R. B., AND L. B. SMITH (1985): "Homeownership Returns, Tenure Choice and Inflation," *American Real Estate and Urban Economics Journal*, 13(4), 343–360.

POTERBA, J. M. (1984): "Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach," *Quarterly Journal of Economics*, 99(4), 729–752.

RAPPAPORT, J. (2007): "Moving to Nice Weather," *Regional Science and Urban Economics*, 37(3), 375–398.

——— (2008): "Consumption Amenities and City Population Density," *Regional Science and Urban Economics*, 38(3), 533–552.

——— (2009): "The Increasing Importance of Quality of Life," *Journal of Economic Geography*, 9(6), 779–804.

RAPPAPORT, J., AND J. D. SACHS (2003): "The United States as a Coastal Nation," *Journal of Economic Growth*, 8(1), 5–46.

ROBACK, J. (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90(6), 1257–1278.

——— (1988): "Wages, Rents, and Amenities: Differences Among Workers and Regions," *Economic Inquiry*, 26(1), 23–41.

ROSEN, S. (1979): *Current Issues in Urban Economics* chap. Wage-based Indexes of Urban Quality of Life, pp. 74–104. Johns Hopkins University Press, Baltimore, MD.

SHAPIRO, J. M. (2006): "Smart Cities: Quality of Life, Productivity and the Growth Effects of Human Capital," *Review of Economics and Statistics*, 88(2), 324–335.

SONG, Y., AND Y. ZENOU (2009): "How Do Differences in Property Taxes within Cities Affect Urban Sprawl?," *Journal of Regional Science*, 49(5), 801–831.

SONSTELIE, J., E. BRUNNER, AND K. ARDON (2000): "For Better or For Worse? School Finance Reform in California," Discussion paper, Public Policy Institute of California.

STEWART, M. B. (1983): "On Least Squares Estimation when the Dependent Variable is Grouped," *Review of Economic Studies*, 50(4), 737–753.

STOVER, M. E., AND C. L. LEVEN (1992): "Methodological Issues in the Determination of the Quality of Life in Urban Areas," *Urban Studies*, 29, 737–754.

WINTERS, J. V. (2009): "Wages and Prices: Are Workers Fully Compensated for Cost of Living Differences?," *Regional Science and Urban Economics*, 39(5), 632–643.

WOOLDRIDGE, J. M. (2006): *Introductory Econometrics: A Modern Approach*. Thomson/South-Western, Mason, OH, 3rd edn.

Part II

Heterodox Economic Perspectives

Chapter 4

Intuitive economics: The role of power and knowledge in the dismal science

4.1 Abstract

Recent disputes over the epistemological nature of social science inquiry, often referred to as "Science Wars", display many of the hallmarks of a much older, largely forgotten quarrel in social science: The *Methodenstreit of Economics* – pitting the Austrian School of Economics against the German Historical School in the 1880s – has largely remained unresolved to date, despite several attempts to overcome it. Yet, curiously, these efforts seem overlooked and outside of the current battlefield of the science wars, where much of the debate is shaped by contemporary political theory. While critical theory, poststructural and postmodern thinking are leaving distinct marks on the sociological research agenda, mainstream economic theory seems impervious to similar developments. Revisiting Edgar Salin's concept of *Anschauliche Theorie* (intuitive theory) and putting it into context with power and knowledge, I argue that such an approach constitutes a vital element for a similar debate in the dismal science. Indeed, a recasting of these ideas as a central pillar within the larger project of socio-economics might be pivotal to developing a new vision in contemporary economic theory.

Keywords: History of economic thought, Methodenstreit, epistemology, power, knowledge, critical theory
JEL Codes: B15, B25, B40

4.2 Introduction

All is not well in the epistemological world of social science inquiry. For most of the post-war half century, the dialogue between qualitative and quantitative discourses in social science has been marred by an increasingly embittered dispute over methodology. Despite the mutual quest for a unifying science, this row has not only deepened many of the cross-disciplinary divides, but it has also created gaping rifts within disciplines. For many fields in social science, the immediate post-World War II period marked indeed a time of "metamorphosis" and departure from disciplinary orthodoxy – a trend that was particularly backed by emerging quantitative methods. While acutely pronounced in economics, this development has re-shaped large parts of its sister disciplines as well, particularly sociology and geography. Across the board, proponents of quantitative science methodology increasingly likened themselves to their natural science counterparts, whereas qualitative methods had become the last bastion of "true social scientists". By the turn of the millennium, these so-called "Science Wars" had almost rendered "qualitative" and "quantitative" analysis into mutually exclusive concepts. With no truce in sight, the stakes for both sides are high as defeat amounts to nothing short of losing relevance with regard to the respective methods of inquiry.

The financial crisis of 2008 has drawn renewed attention to the fault lines of this dispute: The momentous dislocations in the global financial system were not only deemed prime evidence for the failure of modern varieties of capitalism, but they also represented a damning indictment for the science behind the analysis of the global economy. In a recent reply to the Queen's question as to why so few of their guild had foreseen the credit crunch, ten prominent British economists publicly disagreed with the official response from the British Academy (Besley and Hennessy, 2009); rather than a simple combination of wishful thinking and hubris by experts and market participants alike, their dissent puts the blame squarely on the (epistemological) training of economists which had produced "a generation with too many *idiots savants* skilled in technique but innocent of real economic issues"[1]. Indeed, "the preference for mathematical technique over real-world substance diverted many economists from looking at the vital whole, [failing] to reflect upon the drive to specialise in narrow areas of enquiry, to the detriment of any synthetic vision."[2] In a similarly public, but more fervent recent row, German economists at home and abroad have been polarised over the methodological flavour of economic thought that should

[1] Hodgson, Dow, Earl, Foster, Harcourt, Metcalfe, Ormerod, Rosewell, Sawyer, and Tylecote (2009, p.2). Italics in the original.
[2] *Ibid.*

prevail in the curricula at German universities.[3]

These latest disputes over the epistemological nature of social science inquiry display many of the hallmarks of a much older, long forgotten quarrel in the dismal science: The Methodenstreit of Economics – pitting the Austrian School of Economics against the German Historical School in the 1880s – has largely remained unresolved to date despite several attempts to overcome it. Yet, curiously, these efforts seem overlooked and outside of the current battlefield of the science wars, where much of the debate is shaped by contemporary political theory. While critical theory, poststructuralism, and postmodernism are leaving distinct marks on the sociological research agenda, mainstream economic theory seems impervious to similar much-needed developments. Indeed, in this context, the absence of an explicit treatment of the concepts of power and knowledge are particularly surprising if one considers that Bertrand Russell extensively argued that "power to social science is like energy to physics as far as fundamental concepts go".[4] Despite such illustrious support for the integration into the methodological framework of social science, the notion of power and the closely related concept of knowledge are notable only by their absence from contemporary economic analysis. This is especially true of mainstream economics where the paradigm of the neoclassical-Keynesian synthesis has enjoyed an unprecedented hegemony in economic thought.

The current state of affairs, however, is not representative of developments in economic theorising that took place from the late nineteenth century right up until the outbreak of World War II, or at least until the onset of the Great Depression in 1929. Two characteristics of that era are perhaps particularly worth highlighting. First, the intellectual balance of power in social science and economics was equally distributed on both sides of the Atlantic, with a thriving continental European tradition, predominantly lead by German and Austrian thinkers holding their own against a well-established Anglo-American school of thought. Today, however, this balance has shifted dramatically as much of the continental tradition is absorbed into the Anglo-American paradigm. Indeed, one could be forgiven for assuming that the history of the Anglo-American tradition is representative of a broader history of Western economic thought. Second, most leading economic theorists of the day – even the staunchest supporters of the Comtian project of social physics – considered themselves to be social, rather than natural scientists.

Despite being compared to Einstein for his "far-reaching generalization under which Newton's results

[3]German economist Rüdiger Bachmann has kept a comprehensive account of the various contributions to this debate on his website at http://www-personal.umich.edu/r̃udib/methodology.htm. See for example Bent Flyvbjerg's *Making Social Science Matter* (2001) for a more general exposition of the "Science Wars".

[4]Russell (1938, p.ix).

can be subsumed as a special case" (Pigou, 1936, p.3), John Maynard Keynes insisted that economics is essentially a moral science. Indeed, he famously took issue with Lionel Robbins' call for the complete separation of ethics and economics (Atkinson, 2009). Today, such considerations have largely disappeared from the mainstream; the dismal science is deemed to have become as predictive of human action as physics is of nature. Instead, prominent economists deliberate whether their discipline is closer to the practical problem solving in engineering or the development of analytic tools and establishment of theoretical principles in the (natural) sciences.[5] To a large extent, contemporary economics is characterised by a dominance of technique over substance, a development that has given rise to a number of prominent warnings over the last 20 years or so (see Hodgson (2009) for an illustrative treatment of this point).

Revisiting Edgar Salin's concept of Anschauliche Theorie (intuitive theory) and putting it into context with power and knowledge presents a useful starting point for rekindling this debate in the dismal science. Indeed, a recasting of these ideas in the form of "intuitive economics" might constitute a central element within the larger project of socio-economics, aimed at overcoming the lack of vision in contemporary economic theory. The much coveted prize for this challenging venture has many dimensions; most pressingly, it would provide the grounds for a truce in the science wars. Beyond this, it would free the dismal science from its current obsession with method and re-focus policy debates around more pressing questions which have re-emerged in the context of the current wave of capitalist globalisation. Why are economic activity and prosperity spread so unevenly around the globe? Do trade and – more generally – spatial interaction necessarily narrow these differences? What explains the discrepancy between the predictions of theory and what happens in reality? What ever the answers, the challenge for intuitive economics is to rediscover some of the utopian vision of shared by the great economic thinkers. As Golda Meir is famously quoted, "an economist who is no utopist is no good economist" (Salin, 1965, p.227).

The remainder of this paper is organised as follows. Section 4.3 looks at the "Methodenstreit der Nationalökonomie" as the original science war and outlines attempts to resolve the debate, which reveals the first tendency for abstraction of economic theory form the contextual forces of power and knowledge. Section 4.4 looks at how a wedge between theory and reality has lead to a crisis of vision in social science inquiry. Sections 4.5 and 4.6 look at the role of power and knowledge in contemporary political theory,

[5] See e.g. Mankiw (2006) for a recent example.

while contrasting its limited parallel treatment in current economic theory. The notion of intuitive theory, as developed by Salin, provides the analytical starting point for section 4.7 which also attempts to outline the possible elements of a synthesis within socio-economic theory. Section 4.8 provides some tentative reflections on possible new directions for carrying economic thought beyond the science wars.

4.3 Methodenstreit der Nationalökonomie

A *Methodenstreit*, the dispute over methods between different schools of thought, is certainly not a new phenomenon and can be traced back at least to epistemological differences between Plato and Aristotle. In the context of the post-Enlightenment[6] project of social science inquiry, the Methodenstreit between the Historical School and the Austrian School of Economics seems particularly important due to its lasting impact on the methodological debates in contemporary economics and sociology.

Not only did this dispute last over a least three generations from the late nineteenth to the early twentieth century, producing countless intellectual winners, losers and unlikely heirs in its course, it was also shaped by global political developments at the turn of the last century. It was abruptly halted by the Great Depression and seemed resolved, forgotten or no longer relevant during the golden years that followed the trauma of World War II. With the rise of the neoclassical paradigm to unchallenged hegemony from the middle of twentieth century onwards, economic thinkers seemed content to have left tumultuous years of the Methodenstreit behind.

Yet, the fall of real socialism and the collapse of the welfare state dealt an unexpected blow to the neoclassical research programme in which social engineering had made the complete transformation from its normative beginnings as a social science to the hard-and-fast rules of a natural science. As many of the old, methodological fault lines are starting to show again, it seems, therefore, appropriate to return to a brief analysis of the original Methodenstreit and to contemplate the timely relevance of similar developments in other fields of social science.

4.3.1 The original science war

In 1871, not even five years after the first volume of Karl Marx's grand œuvre was published, Austrian economist Carl Menger presented his own interpretation of the political economy, *Principles of Eco-*

[6] I am following Heilbroner (1999) by taking Adam Smith as the first of the modern, post-Enlightenment economic philosophers.

nomics, which would become the intellectual foundations of the Austrian School of Economics. However, it was not until 1883, when the publication of his *Investigations*[7] prompted the *Methodology*[8] by Gustav von Schmoller, a representative of the German Historical School, which marked the beginning of an intense academic debate over the method and epistemological character of economics that should last for several decades.

The Historical School contended that – perhaps in a similar vein to Marx's historical materialism – there was a distinct difference in the fundamental nature of natural phenomena and that of cultural phenomena, the latter of which could only be understood through the interaction of historical processes. According to that view, economics could contribute to the understanding of human action through the study of regularities derived from a historical context.[9] The Austrian School by contrast believed that economics would derive from a basic logical principle and – as socio-economic and political interaction were far too complex to be understood by simple inductive means – a key role of economics would be to develop universally valid theories of human action.

While both schools shared the vision for a universal theory of all social phenomena, the Austrian School – like many of its Anglo-Saxon contemporaries – saw a distinct separation between an economic and a non-economic sphere of human action. The Historical School, on the other hand, strongly emphasised the interdependencies between economic and political developments, deeming it impossible to deduce the complexities of social activities from a single unifying axiom. On a different level, these different positions reflected the classical liberalist conviction[10] of the Austrian School as opposed to the vision of interventionist or welfare-state capitalism by the Historical School. In due course, the Methodenstreit would include thinkers such as Eugen von Böhm-Bahwerk and Ludwig von Mises for the Austrian School and Max Weber and Werner Sombart for the Historical School.

One of the most well-documented attempts to overcome the Methodenstreit is that of Max Weber (1949). He rejected both the descriptive approach of the Historical School and the abstract representation of universal phenomena by suggesting hypothetical "ideal types", intended only serve as heuristic aids for the purpose of understanding a specific case at hand.[11]

---

[7] Menger (1985), "Investigations into the Method of the Social Sciences with Special Reference to Economics".

[8] von Schmoller (1883), "On the Methodology of the Political and Social Sciences".

[9] The epistemological stance of the Historical School also bears a close resemblance to the Foucaultdian method of genealogy, e.g. in Foucault (1995).

[10] Here classical liberalism (or laisser-faire liberalism) is the political tradition that stresses the importance of human rationality, individual property rights, natural rights, constitutional limitations of government, the protection of civil liberties.

[11] Weber's construct of ideal types is frequently interpreted as the positivistic postulate to establish such types, particularly in an

By the late 1930s however – with the original dispute far from being resolved – the uncomfortable association of key thinkers of the Historical School with the Nazi regime[12] and Anglo-Saxon intellectual developments that culminated in Keynes' *General Theory* (1936) brought the Methodenstreit to a premature end. The Austrian School had won by default.

As a consequence, later attempts would no longer enjoy the prominence of earlier ventures to settle the dispute. This includes proposals by Walter Eucken (1950) and Edgar Salin (1944). Eucken's proposal was grounded in a structural notion of the economy as a system that combines the notion of a natural order of things with that created by human action. In contrast to both Weber and Eucken, Salin's proposal identified the increasing mathematical abstraction and the search for universal principles as the underlying causes of what I am calling the "original science wars".

4.3.2 Modern science wars

Undeniably, the significant political crises of the previous century had profound impacts on the conduct of social science, yet their impacts on the research agendas in sociology, political science and economics were very different. While the spectre of fascism and state capitalism in Europe inspired a generation for critical social theorists in what would be called the Frankfurt School, economic theorising grew increasingly void of political questions.

By the middle of the twentieth century, most of mainstream economic theory had become largely apolitical in terms of the questions it addressed and resembled more closely natural sciences, displaying a large degree of mathematisation. As Hands (2006) points out, during the 1950s both economics and the philosophy of science moved from having a variety of competing approaches and research strategies to having a single, almost unanimously accepted mainstream or standard view.

In both disciplines, any notions of power and non-technological knowledge had disappeared from theoretical discourses and, as a result, a new science war was imminent. This time, however, economics had earned the status of a natural science and – from the safety of the vantage point of "true scientists" – contemporary economists are keeping out of harms way in the current debate. At the same time fruitless efforts to provide predictive theory threaten to divorce social science mainstream further from its

institutional context. Others argue that this may indeed be a misinterpretation of Weber's "Idealtypen" which were simply intended as temporary assumptions for analysing specific phenomena. See Swedberg (1999) for a concise overview of this debate.

[12] Sombart's attempt to advise the National Socialists in 1934 had particularly detrimental effects for credibility of the Historical School. Although the party officials rejected his suggestions with deprecating remarks, this tarnished much of his intellectual legacy (Schefold, 2002, p.12).

potential audience, thus pushing it perilously close to the brink of irrelevance.

4.4 Crisis of vision

The Science Wars – both old and new – have at their origins the positivist tradition of science, its proselytic obsession with method and the search for absolute truth. Little sociological and even less economic discourse currently takes place in the public sphere. Instead it manifests itself as "secret writing" in the iterative and disciplinary culture of peer-reviewed journals. This quest for social physics is self-reinforcing by trying to capture the world through quantitative method.[13]

Social science methodology has become ontology through a process of basing itself on scientific concepts that are assumed to "exist outside the text" and that are used to portray cause-and-effect of human interaction in a scientific fashion. Method has by and large erased any authorial presence in sociological or economic writing. It has itself become the main text. This is not a phenomenon that is exclusively limited to empirical social scientists. Even theoretical and qualitative authors are guilty as charged of producing a deliberate science aura compelled by their "science envy". Faced with disciplines that have a penchant for largely being devoid of deliberate authorial choices, and that have been trapped in the "positivistic language games" of academic career building, contemporary scientific social science writing is incapable of solving any intellectual problems using the brute force of technique.

4.4.1 Epistemological problems of economics

Sociologists and economists alike are using "private language" replete with highly charged rhetoric which – by creating the illusions of front-line involvement – serves as a "permanent substitute for experience".[14] The increasing disconnect between social science theory and socio-economic reality constitutes the crisis of vision in the social sciences.

Despite this obvious parallel of increasing "scientism" (i.e. the inappropriate transfer of methods from the natural to the social sciences) in both sociology and economics, the treatment of *power* and *knowledge* marks a fundamental dividing line between the two disciplines. While the original Methodenstreit has certainly contributed to this process in no small measure, prominent sociologists such a Tal-

[13]This is one of the main arguments in Agger's (2000) appeal to return to active *public* sociological practice.
[14]Similar to Agger, Brittan (1983) expresses the concern that academic writing in social sciences has become removed from its readership and audience.

cott Parsons have also cemented this segregation by insisting that institutions, being the embodiments of values, were the proper subject of sociology rather than economics.[15]

As the crisis of vision in sociology may have met its match in the critical writings of contemporary sociologists, similar such developments are few and far between in the dismal science. The criticism by Heilbroner and Milberg (1995) of a disconnect between theory and “the public sphere” in contemporary economic thought marks a notable exception as it bears witness of this state of affairs:

> “The mark of modern-day economics is its extraordinary indifference to this problem. At its peak, the high theorizing of the present period attains a degree of unreality that can be matched only by medieval scholasticism.” (Heilbroner and Milberg, 1995, p.4)

The challenge is thus to rekindle economic debates that are not confined to the ivory towers of academia, but that are derive from “lifeworld grounded critical theory” (Agger, 2006).

4.4.2 Power and knowledge

The relationship between epistemology and power is an important, perhaps even the most significant problem in the contemporary theory of social sciences. For over 70 years, the commonly accepted modes of social science inquiry – separating “power” and “knowledge” (epistêmê or technê)[16] – have been the central subject of attack by the highly abstract writings of critical theorists of the Frankfurt School. More recently, notions such as Flyvbjerg’s “power-augmented phronesis”[17] have been proposed as a way out of the science wars, which are rooted in the wide acceptance of Popper’s positivist paradigm for social theory.

Like in critical social theory, a reintegration of power and knowledge into the framework of economic analysis could help to close that very gap between theory and reality – a gap that is largely attributable to de-politicising and de-historicising of economic thought.

The artificial separation between power and knowledge, between politics and history and – in Lyotard’s (1984) sense – between justice and truth is widely recognised in contemporary political and social

[15] Velthuis (1999) documents Parsons’ objection that institutional economics had a misconceived view on the scope of economics. According to Parsons, institutions as the embodiments of values were the proper subject of sociology rather than economics. This division of labour between economics and sociology – legitimised by Parsons – has shaped much of the interaction between the two disciplines.

[16] Epistêmê ($\epsilon\pi\sigma\tau\acute{\eta}\mu\eta$) is the Greek word most often translated as knowledge, while technê ($\tau\acute{\epsilon}\chi\nu\eta$) is translated as either craft or art. Thus, the relation between the two concepts in ancient philosophy offers an interesting contrast with our own notions about theory (pure knowledge) and (experience-based) practice.

[17] See discussion on power in chapters 7 and 8 of “Making Social Science Matter”(2001).

science. At least since Foucault's (1980) famous contention that "power produces knowledge" and vice versa, the two concepts have become inextricably linked. With this disclaimer and for ease of analysis, however, I will proceed by engaging in a separate treatment of their respective roles in an economic setting.

4.5 Economics and power

As mainstream economics has turned its back on political discourses, power relations have become the blind-spot of economic theory. In the neoclassical framework, power provides the rationale to justify far reaching politically motivated intervention by government. Indeed, in most of economic theory, the notion of power is synonymous with institutions – be it the government or firms – and thus diverges from the concept of the free forces of the market. Government interferes with the market, firms' excess power distorts the market.

4.5.1 Power and institutions

Perhaps surprisingly, social science concepts as diverse as Marxist theories of the state and neoclassical economics share very similar notions of power, either explicit ones in the case of the former or more implicit ones in the latter case. In both instances, power invariably manifests itself in an institutional form and differences only arise with regard to the question of the social group or class within which power is localised. Political theories of pluralism, however, mark a notable exception as power is dispersed equally among a wide variety of diverse interests. In economics, public choice theory and the Tiebout-Tullock hypothesis in particular deal to a limited extend with the dispersion of power outside the institutions of the state.[18]

Be it the state, corporations or civil society, institutional power is best approached from two separate angles. According to Bratsis (2002), for instance, the production of the political community at a national level needs to be distinguished from the formal separation of political from economic power within society (sovereignty vs. social agency). Analytically at least, this seems not very different from the distinction made by neoclassical writers.

Yet, this notion of power seems far too narrow. It does not help to explain the different discursive

[18] See Cebula (1978) for a discussion of this point.

formations at play that take place outside the discrete limits of institutions.[19] The comprehensive understanding of the precise nature and locus of power thus becomes critical – a central point made by Foucault. Indeed, he vehemently argues against such a narrow interpretation of power in a socio-economic context, objecting to the notion that power is limited to a sovereign context and institutional boundaries.[20] Furthermore, power is not an abstract concept that is owned or shared, but it is a pervasive societal process that produces knowledge and truth:

> "[O]ne of the first things that has to be understood is that power isn't localised in the State apparatus and that nothing in society will be changed if the mechanisms of power that function outside, below and alongside the State apparatuses, on a more minute and everyday level, are not also changed." (Foucault, 1980, p.60)

This stands is sharp contrast to the treatment of institutional aspects of power in mainstream economic thought, where the issue is almost exclusively addressed in a distributional context; rather than analysing *how* power works, economists seem more interested in the dichotomy between its equal distribution in perfect competition and its monopolistic distortions of state intervention.

4.5.2 Power and market

In a similar sense to the institutional interpretation of power, authors like von Mises (1996) and Rothbard (2004) – writing in the libertarian tradition of the Austrian School – are positioning "power" and "market" as unequivocal antinomies. Accordingly, the market consists of voluntary transactions between willing parties – firms or individuals – and only the state, or "power", introduces compulsion into human relations, bringing about coerced outcomes that people would not voluntarily have chosen.

In complete contrast, *Pouvoir et économie* by French economist Perroux (1973) develops a notion of power which *is* the market. Like his mentor Léon Walras, Perroux was a Cartesian in method, a socialist in sentiment and an evolutionist in vision. According this notion, general equilibrium thus becomes the interaction of multiple forces which – in accordance to their relative strength – reach a steady state by mutual domination. As power relations are prone to change, equilibrium is inherently instable and unlikely to persist over protracted periods.[21]

[19] This decline of institutions as loci of power is increasingly prevalent in post-industrial, Western societies. Central to this process are the dynamics of sub-politics politics which – centred around the selves – are absorbing the emancipatory role traditionally played by institutions (Beck, Giddens, and Lash, 1994).

[20] See also for example Foucault's (1991) "Governmentality" or his critique of political reason in "Omnes et Singulatim" (1979).

[21] This school of thought, referred to "theory of dominance" has received wide criticism. See Hülsmann (1993) for an overview.

It is not difficult to see the similarities between Perroux's conceptualisation of power and that of his contemporary and countryman Foucault. In both instances, power is interpreted as a disciplinary force that transcends the boundaries of institutions and shapes social conditions. Despite these similarities, Foucault, cautions against the "economism" that he sees present in the conventional analysis of a theory of power.

In the case of contemporary economic theory, for example, power such as sovereignty or market access is taken to be a right that can be possessed like a commodity; in the Marxist conception, power is conceived of as the role it plays in the relations of production and of class domination. In contrast to these "contract-oppression schema" of power, Foucault's self-imposed challenge is to master the art of a *non-economic analysis of power*.

4.5.3 Power and ideology

Both the institutional and the market-based contemplation make it very clear that power cannot be analysed outside of an ideological context. Before proceeding, however, it might be useful to narrow the term ideology somewhat. While ideology is tantamount to a "false consciousness" in its strongest Marxist form, I will rely on a weak definition of the term ideology in the current context. Here, ideology is meant as a set of beliefs, or a specific school of thought, that are concerned with the lasting reproduction of the a specific system of social control and a particular mode of production. Thus, the distribution of power and even the very nature of power itself must be understood through the prism of ideology. In the Marxist tradition, in particular, capitalism is the ideology whose power structures allocate the means of production in the hands of the few by alienating the masses.

In this sense, critical social theory interprets key elements of the capitalist system as authored, ideological text through which power is exercised. Agger (1989), for instance, refers to money as such a modal text which reproduces and stablises the order of capitalism. While such a radical treatment of power may be absent from mainstream economics, critical economists like Kindleberger (1970) and Kirshner (1997) have addressed similar issues from a more traditional perspective.

4.6 Economics and knowledge

In economic theory, the notion of knowledge encompasses many different aspects, most specifically, ranging from the amount of information available about a specific state variable to science and technology in a more narrow sense. In both instances, knowledge is instrumental in the determination of human action and a fundamental organsiational principle of economic activity. Indeed, Prendergast (2010) illustrates how prior to Adam Smith economic progress was largely conceived as being based on the accumulation of knowledge. The development by Turgot and Smith of a concept of capital that included other factors contributing to development led their followers to focus on capital to the neglect of the independent role of knowledge. With this in mind, these two forms of knowledge will be examined in more detail.

Apart from the problems of "scientism" addressed in section 4.4, the so-called "knowledge problem" has been central to the economic thought of proponents of the Austrian School, such as Friedrich von Hayeck, who was a student of Ludwig von Mises. von Hayek's (1937) poineering paper "*Economics and Knowledge*" defines the necessarily incomplete and dispersed knowledge of economic agents as *the* departing point of economic research. In his view, it was never the primary problem of economic theory to determine the mathematical conditions of general equilibrium, but rather how the voluntary coordination of individual plans is achieved.

4.6.1 Praxeology and knowledge

Central to this interpretation of knowledge are von Mises' beliefs that the nature of economic activity is just a special sub-set of all human action, which itself is fundamentally governed by values. In distinct contrast to the natural sciences it is not based on observations which can be refuted on the basis of empirical investigations. Instead, this set of *a priori laws* can only be discovered through deductive reasoning. This is the role of general social science, which – in order to distinguish it from Comtian sociology – von Mises referres to as *praxeology*.

Praxeological foundations are laid out by von Mises's axiom of human action or purposeful behavior. Society is a product of the human urge to remove uneasiness and dissatisfaction as far as possible; it is not a product of social classes, political hierarchies, and various other synthetic structures. In this decidedly Kantean view of social activity, economics is merely a subdiscipline of praxeology dealing with the laws of human action in a system of private property of the means of production.

According to thinkers of the Austrian School, praxeology *is* epistemology. As such, it is diametrically opposed to Marx's historical materialism which explains how and why historical events occur through the bourgeois-versus-proletariat class struggle.

Through the *principle of immanence*, the "explanation of every event as repetition", praxeology links directly to the critiques of human (in)agency that characterise the writings of Horkheimer and Adorno (2002). They would undoubtedly view praxeology as no less mythical and problematic than mainstream sociology. In their views, enlightenment reveals itself as operating in the same way as the myth itself. Through abstraction which represents the instrument of enlightenment, it stands "in the same relationship to its objects as fate, whose concept it eradicates: as liquidation" (Horkheimer and Adorno, 2002, p.9). Enlightenment and the immanence of positivism liberates man from the mythical fear of the unknown ("the outside"), yet the universal taboo that there is no outside thus becomes "mythical fear radicalised" (*Ibid.*, p.11)

Indeed, Horkheimer and Adorno view enlightenment as totalitarian as any system can be, not – as is often criticised – because of its analytical method, but because it assumes "that the trial is prejudged"(*Ibid.*, p.19). In other words, while praxeology is critical towards positivist sociology on epistemological grounds, Horkheimer and Adorno would declare both projects are flawed, but on different grounds. Enlightenment has created *homo oeconomicus* as a logical subject that forms the reference point for reason, preoccupied only with his own self-preservation, yet ultimately incapable of any agency.

4.6.2 The production of knowledge

One of the defining characteristics of the current era of globalisation is the increasing emphasis on the notion of the "knowledge economy", an expression of the commodification of knowledge. While the first wave of globalisation at the end of the nineteenth and early twentieth century was largely driven by the process of industrialisation, knowledge is at the centre of several processes that operate within contemporary globalisation. Put differently, knowledge has become the performitivity principle of the modern age, which – according to postmodern theorists like Lyotard (1984) – has led to a crisis of scientific knowledge. I will return to this point shortly.

A simple description of the production process allows us to distinguish between two elements of knowledge. Using economic theory, we can define a given output as the specific combination of a set of scarce input factors of production, capital and labour. In its most simple form thus described, knowledge enters the production process – or its neoclassical analytical equivalent, the production function – in two separate ways: either as a specialised factor of production in the form of *human capital*, or as the *technology* required to achieve the "specific combination" of factor inputs. In the compelling shorthand of economic analysis, knowledge can thus be reduced to a mere set of parameters. In terms of the most popular neoclassical production function, the Cobb-Douglas function, this is expressed as $Y = AK^{\alpha}L^{\beta}$

where the parameters A, α and β depend on technology, i.e. knowledge.

In this setting, one of the most convenient properties of this knowledge-based production function is its scale-invariance. Despite the many items of disagreement between the "new economic geography" (NEG) and economic geography "proper" (see chapter 1), both share the necessity for some form of aggregation of human capital across agents and scale in order to reduce analytical complexity. However, the "Cambridge capital controversy" (e.g. Harcourt, 1996; Cohen and Harcourt, 2003) has highlighted that the spatial configuration of economic activity may in reality be very sensitive to aggregation when capital is not homogeneous. Indeed, the Cambridge debate showed conclusively that the aggregation of capital, so as to obtain a surrogate production function according to Samuelson, is not possible in general, with critical implications also for other variants of neoclassical theory (e.g. Robinson, 1971).

Aside from these issues, in many instances knowledge is both the human capital *and* the technology required in the production process. Furthermore, in the context of a global production process, knowledge is defined in a positivistic, Western sense. The global division of labour therefore either excludes other forms of knowledge or it commoditises them. For instance, intellectual property rights instrumentalise indigenous knowledge and commodetise them for the production process. In this sense, knowledge is also a critical element of "academic imperialism". This is a practice that further entrenches the economic dependency of developing countries at the periphery by imposing scientific structures prevalent at the core.

It is this performative production of knowledge, however, that has eroded its legitimation and has caused a "scientific crisis" in late capitalist society. Universities as the production facilities of capitalist knowledge are at the centre of this crisis, a phenomenon that receiving increasing attention from critical theorists. Luke (2005), for example, examines and characterizes the precise condition of contemporary science discourse and research at American Universities. Unsurprisingly, perhaps, his verdict is grim – merely a reflection of the contextual vacuum the academic disciplines are currently finding themselves in. Academics are trading in pedagogy for *performativity* as the modern research university encourages the commoditisation and monetization of knowledge. Instead of seeking "truth, progress or freedom", universities are putting knowledge to work and corporations take control over intellectual property and "journal science" becomes the main transmission mechanism of knowledge. Luke argues that globalization is shaping a perception of "legitimate knowledge" whose forces reshape universities according to the needs of the knowledge market, using "interdisciplinary structures" to break up general education.

The cumulation of knowledge as a means of progress has given way to a subordination of knowledge to the technological means of production in order to "reinforce reality" – thus reversing the relationship between knowledge and technology. Lyotard describes this as the legitimation crisis of the "speculative narrative" of knowledge:

> "The State and/or company must abandon the idealist and humanist narratives of legitimisation in order to justify the new goal: in the discourse of today's financial backers of research, the only credible goal is power. Scientists, technicians, and instruments are purchased not to find truth, but to augment power." (Lyotard, 1984, p.46)

There are, of course, distinct parallels to the Foucaultian case where power is producing knowledge. Most interestingly, perhaps, is the fact that this new mode of knowledge production is not only discussed by continental philosophers, but also constitutes a widely recognised phenomenon in the business management literature. There, however, it is simply described as *Mode 2* knowledge production.[22]

4.6.3 New Science and New Technology

In contrast to the instrumental rationality of knowledge in economic thought and its much debated crisis of legitimation in postmodern theory, critical theorists like Marcuse deem the critique of science and technical rationality as the first step towards overcoming domination.[23] In this context, it is not the accumulation of knowledge that is the emancipatory force, but overcoming technical rationality that holds the key for social change. Agger (1976) describes Marcuse's concept of "new sensibility" – a wake-up call for human agency – as new science which is

> " [...] a cognition freed from positivist fact-fetishism and a dualist theory of knowledge. The new science is a mode of thought and imagination which engages in speculation for its own sake, a from of nonalienating work." (Agger, 1976, p.161)

Like Marx, Marcuse views the social division of labour, maintained by a technocratic professionalism, as the root cause of domination, in particular in the context of the alienation of labour. Marcuse puts technical rationality at the centre of domination, "surplus repression" and one-dimensionality. Technical rationality permits technical domination as the highest form of exploitation and needs to be abandoned in the quest for a "new science". However, this critique of science and technology stands in strong

[22] This expression was first coined by Novotny et al. (1994) who use define mode 1 knowledge production as the traditional, academic and investigator-initiated and discipline-based production, while mode 2 is problem-focused and interdisciplinary and involves multidisciplinary teams which work on specific problems in the "real world", i.e. creating reality.

[23] In *One-Dimensional Man*, Marcuse (1991) describes technical rationality as a totalizing ideology creates truth and reality.

contrast to Jürgen Habermas's view that the instrumental rationality of science can be salvaged without undermining the critical role of social philosophy.

4.7 Intuitive economics

The previous sections have highlighted the many direct consequences of a mainstream tradition of economic thought that abstracts its analysis from a political and historical context and from the force field of power and knowledge. The purpose of this section is to plead the case for a re-discovery of the political economy as the centre of social science inquiry.

The epistemological ideals of Werner Sombart (1933) and Edgar Salin (1944, 1965) – who worked extensively out of Sombart's work – are of particular relevance to this venture. Conceptually, it relies on "Anschauliche Theorie"[24], an idea that was first mentioned by Sombart and then further developed by Salin. Most importantly, perhaps, intuitive theory does not only aim at an intellectual retracing of the causal relationships that govern economic activity intellectually, but it constitutes a critical theory of economic, social and cultural change.

4.7.1 Economics as a social science

As indicated, the origins of intuitive theory in economics are found in Sombart. He was not only a prolific writer – probably best remembered for his monumental, three volume tome entitled *"Der Moderne Kapitalismus"*(1927), which traces the rise of capitalism over some 2,350 pages–, but he was also one of the most vociferous opponents of the exclusive reliance on natural science methods in economics. His efforts to redress the balance against the rise of scientism lead him to distinguish the *"verstehende"* (understanding) economics from the other two types, the *"ordnende"* (ordering) and *"richtende"* (judging) which are present in classical economic thought.

Opposed to the method of natural sciences, Sombart contends that there is a method appropriate to social science inquiry that deals with culture – the method of the *Kultur*- or the *Geisteswissenschaften*.[25] However, Sombart is well aware of the complications that such an undertaking entails:

> "It is the certainly unenviable fate of our entre discipline that it is incapable of gaining anything other

[24] Ger. intuitive or descriptive theory; hereafter translated as intuitive theory

[25] Ger. Cultural Sciences or Humanities.

than intuition in the realm of cultural knowledge (or science)." (Sombart cited in Schefold (1992, p.317), my translation from German.)

While social science inquiry may engage in some of the same quantitative exercises as natural sciences, their ultimate task is to enable a deeper understanding of the processes of economic activities.

4.7.2 Edgar Salin's "Anschauliche Theorie"

Like Sombart and the brothers Weber, Salin was an economist with sociological erudition. While a student of Alfred Weber's, he reportedly admired Max Weber's personality, but remained highly critical of Weber's reliance on abstract concepts to explain social phenomena.[26] Salin argued that the question to what extent economics was to be viewed a social science, or even part of the humanities, principally depended on the scientific questions that the field addressed – not the methods it deployed. In the context of the modern science wars, however, the emphasis almost exclusively rests on method. If the predominance of natural science method in contemporary economics amounts to more than "emulation as sincerest form of flattery" or simply is a consequence of "science envy", its origins are rooted in classical economic theory.

In the philosophy of the great classical economists, the dependence of economics on natural science ideals is motivated by the desire to interpret the capitalist mode of production as a process which was to be liberated from political interference. Indeed, this interpretation reached an early paradigmatic peak with the marginalist or neoclassical revolution. Neoclassical economics has since become a new standard of economic thought – a standard which has completely done away with any hermeneutics in order to understand cultural embedding of economic activity. In the neoclassical system – still the reigning paradigm in economics today – even consumers' preference sets, in the form a utility function, are assumed as determined. A reflective interpretation on historical conditions and circumstances is no longer viewed as key element of economic analysis.

Intuitive or descriptive theory in the sense of Sombart, and particularly as developed further by Salin, begs to differ. Salin's work was guided by the desire to establish intuitive theory as a viable alternative to the doctrine of the Chicago School which depends exclusively on neoclassical rationality. While still relying on rational theory, intuitive theory also "incorporates sociological and psychological moments to

[26] Schefold (2004, p.3), a student of Salin's, reports an illustrative exchange between Weber and Salin over Theodor Mommsen's Nobel prizewinning work on the history of the Roman Empire. Asked what he thought of it, "Weber, upon this, very loud: 'That is no science!' Salin replicated: 'Then I don't know how your science could serve what is alive and why it should be of interest for us'."

deepen the understanding of Capitalism" (Schefold, 1992, p.304), particularly with regard to interdependencies in the investment process or the economic cycle.

It is not difficult to see the parallels to concepts developed by some of his contemporaries – intuitive concepts such as Schumpeter's business cycles or Keynes' "animal spirits" which have survived the sandblasting of the neoclassical synthesis after the Second World War. Although admiring Keynes for his writing skills and Schumpeter for his sociology of the entrepreneur, Salin showed contempt for formalistic model building. He vehemently criticised abstract models as "partial knowledge" – not devoid of truth, but erroneous because of the generalisations usually based on it (Schefold, 2004, p.7).

4.7.3 Elements of a synthesis

Intuitive theory is inspiration, not romantic irrationality as Salin's many critics – including Weber – were quick to condemn. Far from clinging to a nostalgic backward gaze, Salin hoped to resuscitate those forces in scientific inquiry that were being suffocated by the preponderance of rationalism. Indeed, Salin speaks of Weber's attempts to establish objective, value-free science as

> "[...] the tragic courage of a lost generation [which] draws the conclusion from the palsy of the old values, ideals and religions that there are no Gods, no set ranks and no all-binding measures, thus becoming the propagators of the last level of de-selfed and de-spirited work." (Salin (1932) cited in Schefold (1992, p.317), my translation from German.)

Intuitive theory does not reject rationality outright – it keeps a rational core that explains the functioning of a particular system while relying on descriptive components to explain the totality of economic activities. With a historical focus similar to that of the work of Marx, Weber and Schumpeter, intuitive theory depends on context and aims to marry economic analysis with sociological or cultural insights. Salin thus proposes intuitive theory as encompassing rational theory in the sense of either classical or neoclassical theory of value, while capable of describing the consequences of other forms of motivation.

In other words, intuitive theory integrates both power and knowledge in a way that most of mainstream economics has come to abstract from them. Beyond the "partial cognition" of the latter, the former aims high to visualise the "*Gestalt* of capitalism".[27] In intuitive theory, capitalism is a totality that neither resulted from a concentration of experience nor from hypothetical abstractions of some logical principle alone – neither empirical phenomena nor constructed ideal types are capable of exhausting the

[27] This terminology originates with Goethe and forms an important analytical lens throughout Salin's work.

term ‘capitalism’. They just constitute partial cognition, whereas intuitive theory relies on a synthesis of these elements to achieve “total cognition” (Gesammterkenntnis).

Salin was conservative liberal who was critically influenced by Marx. He was well aware of the fundamental contradictions that straddle the capitalistic process of industrial concentration when he writes

> “[n]ow the elements of creation have been found and conquered by humanity – by a weak generation which disposes only of a strong intelligence but which is neither guided by faith nor by a feeling for responsibility. Technique as an instrument thus has been transformed in technique as a sceptre or – in a whip for slaves.” (Salin (1957) cited in Schefold (2004, p.6))

Despite his occasionally pessimistic tone in the context of cultural development, Salin tried to sythesise the combined effects of technological and economic progress. Relying on context and value-rationality allows intuitive theory to be strong where natural science is weak. In this sense, Salin’s intuitive theory is precisely what – almost over half a century later – Flyvbjerg (2001) re-invents and re-labels as “phronetic social science”. His plea for induction and the importance of case studies as a scientific method *is* intuitive theory. More “strategic sampling” and more “good narrative” which resists the temptations of generalization *is* intuitive theory. While Flyvbjerg relies on the analyis of Foucault to provide the tools to understanding power in relation to phronesis, Salin is directly grounded in the orignial traditions of Marx, Weber and Sombart.

Be it intuitive theory or phronetic methods, the political economy is the focal point of social science inquiry for both. In this context, economists today would do well to heed Salin’s words:

> “All economic science is – in its intrinsic nature and by its objectives – a political science [...] and will thus remain, from the very beginnings into the future, a study of the political economy.” (Salin (1965, p.16), my translation from German.)

Indeed, focusing on values steeped in situational ethics and contextualism and placing power at the core of the analysis, Flyvbjerg is paraphrasing Salin when he proposes a de-centred method of phronetic social inquiry. Rather than fretting over parametric versus non-parametric methods or loosing sight when adding the n-th equation to unwieldy dynamic stochastic general equilibrium models, economists should start asking again *“where are we going?”*, *“is it desirable?”* and *“what should be done?”*.

4.8 Outlook

The absence of various notions of power and knowledge contemporary economic analysis have gone hand in hand with the increasing mathematical formalisation of and reliance on method in the dismal science. Yet, the analysis of power may not constitute an explicit element of the debate, contemporary mainstream economic thought nonetheless makes important assumptions about the relationship between power and economic activity. These assumptions are perhaps even more important than those about the efficiency of free markets. Indeed, on the relationship between economic freedom and political freedom, Milton Friedman argues

> "[...] the kind of economic organization that provides economic freedom directly, namely, competitive capitalism, also promotes political freedom because it separates economic power from political power and in this way enables the one to offset the other. Historical evidence speaks with a single voice on the relation between political freedom and a free market. I know of no example in time or place of a society that has been marked by a large measure of political freedom, and that has not also used something comparable to a free market to organize the bulk of economic activity. [...] By removing the organization of economic activity from the control of political authority, the market eliminates this source of coercive power. It enables economic strength to be a check to political power rather than a reinforcement." (Friedman, 1962, pp.15–17)

Beyond the destructive forces of the science wars, intuitive theory provides a viable alternative to overcome the current crisis of vision of contemporary economic thought that is mired in method. Intuitive theory provides a much neglected analytical framework that integrates power and knowledge, while not completely jettisoning the cultural heritage of rational economic theory.

While their Marxist legacies have inspired a notable array of critical methodologies in sociology or political science, such projects are in economics. Mandel's *Late Capitalism* (1978) marks an important exception. In there are close similarities to the work of Sombart and Salin, in more than just the name.[28] Although the treatment of epistemology "foreshadows the book's most critical weakness" to some (Hunt, 1977, p.336), a more generous interpretation permits to recognise elements that are very closely related to Salin's intuitive approach.

[28] Cf. Sombart's *Der Moderne Kapitalismus* and Salin's essay entitled *Hochkapitalismus*. Curiously, this parallel seems to have escaped Mandel himself. Judging by references in *Late Capitalism*, his awareness of a German tradition of economic thought other than Marx seems decidedly limited.

Whereas Mandel seems hellbent on equating dialectical method of Marxism as the only "true science", Salin is less radical and insists on the complementarity of both rationalism and empiricism. Unfortunately, the alienating tone of Mandel's dialectical jargon drowns out many of his interesting hypothesis that could be of great importance to understanding contemporary capitalism. Both Marx's historical materialism and Salin's intuitive theory recognize that science is necessary precisely because "essence and appearance" never directly coincide. Both do not ascribe to science only the task of the discovery of the essence of economic relations which are obscured by their superficial appearances. But both regard social science as "the explanation of these appearances themselves, in other words as the discovery of the intermediate links, or meditations, which enable essence and appearance to be reintegrated in a unity once again." (Mandel, 1978, p.22).

Where this integration fails to occur, theory is reduced to the partial cognition of abstract models which miss the totality of empirical reality. Revealing some of his Marxist heritage, Salin rightly argues that the intellectual reproduction of reality must remain in constant contact with the actual movement of history. Only embedding the rational analysis of economic activities within the total cognition of intuitive theory will reveal the true *Gestalt* of capitalism – Salin's most ambitious goal.

4.8.1 Globalisation as common origins

What are the implications in the current context? One immediate historical parallel springs to mind, namely a possible connection between the science wars and the challenges of globalisation. Indeed, the Methodenstreit coincides with first wave of globalisation 1820–1914, whereas origins of current science war are to be found at beginning of second wave in the late 1950s. One of the defining characteristics of the current era of globalisation is the increasing emphasis on the notion of the "knowledge economy". While the first wave of globalisation at the end of the 19^{th} and early 20^{th} century was largely driven by the process of industrialisation, knowledge is at the centre of several processes that operate within contemporary globalisation.[29]

Undeniably, the process of capitalist globalisation is having profound impacts on national economies and – through related, though distinctly separate processes – also on the nation-state. What does this severance of the economy and state imply for the theory of economic thought and policy making. Mandel's *Late Capitalism* may offer some elements of a synthesis – in strictly Marxist terms – that might help

[29] See Baldwin and Martin (1999) for an extensive comparison of these two waves of capitalist globalisation.

explain the underlying causes of the current wave of globalisation. While it is difficult to agree with his assertion that "the long post-war wave of rapid growth [...] to both non-Marxist and Marxist economist by surprise" (Mandel, 1978, p.7), the process of globalisation is posing unprecedented challenges to both policy makers and economic agents alike. Intuitive theory, however, may yet present the most versatile tool to cope with the complexities of globalisation comprehensively.

4.8.2 Late Capitalism or Economic Renaissance?

The heart of Mandel's analysis is his extensions of Kondratieff's waves of economic activity – a variation on Schumpeter's hypothesis in *Business Cycles* (1939). As noted earlier, the use of such intuitive phenomena to capture economic activity are also critical to Salin's postulate. In contrast, however, Mandel's verdict bears many elements of much maligned Marxist determinism as he predicts that the postwar boom would "be followed by another long wave of increasing social and economic crisis for world capitalism, characterised by a far lower rate of economic growth" (Mandel, 1978, p.119).

Nonetheless, recent evidence suggests that globalisation may indeed have exacerbated wage inequality in developed countries, while – perhaps paradoxically – the impact of globalisation on both income and earnings inequality in less-developed countries has been negligible. (Dreher and Gaston, 2006). Furthermore, the globalisation-induced metamorphosis of the national economy has been widely accepted as an empirical fact. The interpretations of the possible consequences, however, remain subject to a variety of different interpretations by contemporary political scientists, sociologists and economists alike.

Beck (2005), for example, argues for a different, receding influence of the state in second modernity by augmenting the concept of "sub-politics" with the "politics of side-effects". The politics of side-effects is "domination by nobody" because it "[...] spares those who benefit from it the trouble of having to organise and legitimise themselves as political actors, without necessarily forfeiting any political power in the process". (Beck, 2005, p.117). Indeed, rather than the power of the state it is the counter-power of global civil society that forms the basis of this "delegitimisation of domination", the key trend of transition from the first to the second modernity. This change in the locus of political power – the power of self-legitimisation – leads to knowledge and understanding of what is good and replaces democratic legitimisation. "Cosmopolitan democracy", so Beck, is a mere fig leaf used to conceal the undemocratic, moral, metaphysical self-justification of *ademocracy* (not anti-democracy) in the diverse cosmopolitan society.[30]

Static methods as they are currently wide-spread in the dismal science are inadequate tools to analyse such transformations of power structures and knowledge. Nonetheless, Hall and Gingerich (2009) demonstrate some statistical analysis of that the core contentions of the 'varieties of capitalism' perspective on comparative capitalism can yield important insights. Whether the state intervenes on behalf of capitalism or globalisation constitutes an economic renaissance that is marked by the demise of this very state, economists need to re-tool in order to contribute meaningfully to the varied challenges of our current times. As in any epoch that faces transition, a purposeful utopian vision rather than a melancholic backward gaze holds the key to manage change successfully. During the last years of the Second World War – as the indescribable terror and turmoil of human conflict had brought death and destruction over much of Continental Europe – the German economist and founding director of the German Institute for Economic Research (DIW) Ernst Wageman (1943) summarises this feeling most succinctly:

> "It rises from the deepest nature of our time that we do not ask in peculiar reflection so much, whether our social, political, economic acting stands in agreement with the requirements of old tradition, but whether it is fruitful, whether it pleases our people, our new generation and later generations." (Wageman (1943, p.314), my translation from German.)

Whatever the stage or variety of capitalism, adequate social science needs to reflect that the interplay between culture and economy are not static and thus cannot be analysed using static methods. Therein lie the true origins of the crisis of vision in contemporary economics. Indeed, Beckert and Streeck (2008) suggest that "the most promising approach is close cooperation between the scholarly traditions of political economy and economic sociology, with the former standing to benefit from a more explicit microfoundation in a sociological theory of action and the latter from more systematic consideration of politics and the state." In a similar vein, intuitive economics is a much needed return to utopian thinking in the dismal science.

4.A Acknowledgements

I am indebted to Patricia Mooney Nickel and Peter Schaeffer for insights, discussion and vital suggestions on an earlier draft which was circulated under the title "Avoiding a third crisis of economic theory?

[30] *Ibid.*, p.297. Beck's notion of cosmopolitan democracy bears a close resemblance to Hardt and Negri's network society (2000).

Intuitive economics and the role of power and knowledge reconsidered". Inputs from participants at the 2007 Society for the Advancement of Socio-Economics are also gratefully acknowledged.

4.B References

AGGER, B. (1976): "Marcuse & Habermas on New Science," *Polity*, 9(2), 151–181.

——— (1989): *Fast Capitalism: A Critical Theory of Significance*chap. Fast Capitalism?, pp. 15–22. University of Illinois Press, Urbana and Chicago.

——— (2000): *Public Sociology: From Social Facts to Literary Acts*, New Social Formations. Rowman & Littlefield Publishers, Inc., Lanham, MA., first edn.

——— (2006): *Critical Social Theories: An Introduction* chap. The Politics of Grand Narratives I, pp. 34–55. Paradigm Publishers, 2nd edn.

ARONOWITZ, S., AND P. BRATSIS (eds.) (2002): *Paradigm Lost: State Theory Reconsidered*. University of Minnesota Press, Minneapolis.

ATKINSON, A. B. (2009): "Economics as a Moral Science," *Economica*, 76(s1), 79–804.

BALDWIN, R. E., AND P. MARTIN (1999): *Globalisation and Labour*chap. Two Waves of Globalisation: Superficial Similarities, Fundamental Differences. J.C.B. Mohr for Kiel Institute of World Economics, Tübingen.

BECK, U. (2005): *Power in the Global Age: A New Global Political Economy*. Polity Press, Cambridge, first published as "Macht und Gegenmacht im globalen Zeitalter: Neue weltpolitische Ökonomie", Suhrkamp Verlag, 2002.

BECK, U., A. GIDDENS, AND S. LASH (1994): *Reflexive Modernization: Politics, Tradition and Aesthetics in the Modern Social Order*. Stanford University Press, Standford.

BECKERT, J., AND W. STREECK (2008): "Economic Sociology and Political Economy: A Programmatic Perspective," MPIfG Working Paper No. 08/4, Max Planck Institute for the Study of Societies, Cologne.

BESLEY, T., AND P. HENNESSY (2009): "Reply to HRH the Queen," Letter, British Academy.

BLAUG, M. (1997): *Economic Theory in Retrospect*. Cambridge University Press, Cambridge, U.K., fifth edn.

——— (1998): "The Problems with Formalism," *Challenge*, 1, 42–48.

BRITTAN, A. (1983): "Sociology as a Private Language," *Poetics Today*, 4(3), 581–594.

CEBULA, R. J. (1978): "An Empirical Note on the Tiebout-Tullock Hypothesis," *Quarterly Journal of Economics*, 92(4), 705–711.

COHEN, A. J., AND G. C. HARCOURT (2003): "Retrospectives: Whatever Happened to the Cambridge Capital Theory Controversies?," *Journal of Economic Perspectives*, 17(1), 199–214.

DREHER, A., AND N. GASTON (2006): "Has Globalisation Increased Inequality?," Working Paper 140, Swiss Institute for Business Cycle Research, Swiss Federal Institute of Technology Zurich.

EUCKEN, W. (1950): *Die Grundlagen der Nationalökonomie*. Springer Verlag, Berlin, Göttingen, Heidelberg, 6 edn.

FLYVBJERG, B. (2001): *Making Social Science Matter: Why social inquiry fails and how it can succeed again*. Cambridge University Press, Cambridge, first edn.

FOUCAULT, M. (1979): "Omnes et Singulatim: Towards a Criticism of 'Political Reason'," in *The Tanner Lectures on Human Values*, pp. 225–254. Stanford University.

——— (1980): *Power/Knowledge: Selected Interviews & Other Writings 1972–1977*. Pantheon Books.

——— (1991): *The Foucault Effect: Studies in Governmentality*chap. Governmentality, pp. 87–104. University of Chicago Press, Chicago, first edn.

——— (1995): *Discipline & Punish: The Birth of the Prison*. Vintage Books, New York City, reprint edn.

FRIEDMAN, M. (1962): *Capitalism and Freedom*. University of Chicago Press, Chicago, first edn.

HALL, P. A., AND D. W. GINGERICH (2009): "Varieties of Capitalism and Institutional Complementarities in the Political Economy: An Empirical Analysis," *British Journal of Political Science*, 39(3), 449–482.

HANDS, D. W. (2006): "A Tale of Two Mainstreams: Economics and Philosophy of Natural Science in the mid-20th Century," Presidential address, History of Economics Society.

HARCOURT, G. C. (1996): "Critiques and Alternatives: Reflections on Some Recent (and not so Recent) Controversies," *Journal of Post-Keynesian Economics*, 19(2), 171–180.

HARDT, M., AND A. NEGRI (2000): *Empire*. Harvard University Press, Cambridge, Mass.

HEILBRONER, R. L. (1999): *The Worldly Philosophers: The Lives, Times And Ideas Of The Great Economic Thinkers*. Touchstone, seventh edn.

HEILBRONER, R. L., AND W. S. MILBERG (1995): *The Crisis of Vision in Modern Economic Thought*. Cambridge University Press, Cambridge, UK, first edn.

HODGSON, G. M. (2009): "The Great Crash of 2008 and the Reform of Economics," *Cambridge Journal of Economics*, 33(6), 1205–1221.

HODGSON, G. M., S. C. DOW, P. E. EARL, J. FOSTER, G. C. HARCOURT, J. S. METCALFE, P. ORMEROD, B. ROSEWELL, M. C. SAWYER, AND A. TYLECOTE (2009): "Reply to HRH the Queen: Why Few Economists Had Foreseen The Credit Crunch," Discussion paper, The Business School, University of Hertfordshire.

HORKHEIMER, M., AND T. W. ADORNO (2002): *Dialectic of Enlightenment: Phlilosophical Fragments*. Stanford University Press, Stanford.

HUERTA DE SOTO, J. (1998): "The Ongoing Methodenstreit of the Austrian School," *Journal des Economistes et des Etudes Humaines*, 8(1), 75–113.

HÜLSMANN, J. G. (1993): *Kritik der Dominanztheorie: Zur machttheoretischen Wirtschaftsanalyse Francois Perroux' und zur Bedeutung von auf sie zurückgehenden wirtschaftspolitischen Vorschlägen*. Rita Fischer, Frankfurt a.M.

HUNT, E. K. (1977): "Late Capitalism – A Review," *Contemporary Sociology*, 6(3), 335–337.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest and Money*. Macmillan, London, 1st edn.

KINDLEBERGER, C. P. (1970): *Power and Money: The Economics of International Politics and the Politics of International Economics*. Basic Books, New York.

KIRSHNER, J. (1997): *Currency and Coercion: The Political Economy of International Monetary Power*. Princeton University Press, Princeton, paperback edn.

LUKE, T. W. (2005): "From Pedagogy to Performativity: The Crises of Research Universities, Intellectuals and Scholarly Communication," *Telos*, 1(131), 13–32.

LYOTARD, J.-F. (1984): *The Postmodern Condition: A Report on Knowledge*, vol. 10 of *Theory and History of Literature*. University of Minnesota Press, Minneapolis, 1 edn.

MANDEL, E. (1978): *Late Capitalism*. Verso, London, verso edition edn., first published as "Der Spätkapitalismus", Suhrkamp Verlag, 1972.

MANKIW, N. G. (2006): "The Macroeconomist as Scientist and Engineer," *NBER Working Paper No.12349*.

MARCUSE, H. (1991): *One-Dimensional Man: Studies in the Ideology of Advanced Industrialised Society*. Beacon Press, Boston, second edn.

MENGER, C. (1985): *Investigations into the Method of the Social Sciences with Special Reference to Economics*. New York University Press, New York.

NOWOTNY, H., C. LIMOGES, M. TROW, S. SCHWARTZMAN, P. SCOTT, AND M. GIBBONS (1994): *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. Sage Publications Inc, first edn.

PERROUX, F. (1973): *Pouvoir et économie*. Bordas, Paris.

PIGOU, A. C. (1936): "Mr J.M. Keynes' General Theory of Employment, Interest and Money," *Economica*, 3(10), 115–132.

PRENDERGAST, R. (2010): "Accumulation of Knowledge and Accumulation of Capital in Early ŚTheoriesŠ of Growth and Development," *Cambridge Journal of Economics*, 34(3), 413–431.

ROBINSON, J. (1971): "The Measure of Capital: The End of the Controversy:," *Economic Journal*, 81(323), 579–602.

ROGIN, L. (1933): "Werner Sombart and the "Natural Science Method" in Economics," *Journal of Political Economy*, 41(2), 222–236.

ROTHBARD, M. N. (2004): *Man, Economy, and State*chap. Power and Market, pp. 1047–1367. Mises Institute, scholar edn.

RUSSELL, B. (1938): *Power: A New Social Analysis*. Routledge, New York.

SALIN, E. (1944): *Geschichte der Volkswirschaftslehre*. Francke, Bern, 4th edn.

——— (1965): *Politische Ökonomie: Geschichte der wirtschaftspolitischen Ideen von Platon bis zur Gegenwart*. Mohr, Tübingen, 5th edn.

SCHEFOLD, B. (1992): "Nationalökonomie als Geisteswissenschaft – Edgar Salins Konzept einer Anschaulichen Theorie," *List Froum für Wirtschafts- und Finanzpolitik*, 18(4), 303–324.

——— (2002): *The Future of the History of Economics*chap. Reflections on the Past and Current State of the History of Economic Thought in Germany, pp. 125–136. Duke University Press, Durham and London.

——— (2004): "Edgar Salin and his Concept of 'Anschauliche Theorie' (Intuitive Theory) During the Interwar Period," *Annals of the Society for the History of Economic Thought*, 46, 1–16.

SCHUMPETER, J. A. (1939): *Business Cycles: A Theoretical, Historical and Statistical Anaysis of the Capitalist Process*, vol. I & II. McGraw-Hill, New York, first edn.

SOMBART, W. (1927): *Der Moderne Kapitalismus*. Dunkler und Humblot, München und Leipzig.

SWEDBERG, R. (1999): "Max Weber as an Economist and as a Sociologist," *American Journal of Economics and Sociology*, 58(4), 561–582.

VELTHUIS, O. (1999): "The Changing Relationship between Economic Sociology and Institutional Economics," *American Journal of Economcis and Sociology*, 58(4), 629–649.

VON HAYEK, F. A. (1937): "Economics and Knowledge," *Economica*, 4(13), 33–54.

VON MISES, L. (1978): *Epistemological Problems of Economics*chap. Introduction to the Third Edition, pp. 1–52. Mises Institute, Auburn, Alabama, third edn.

——— (1996): *Human Action: A Treatise on Economics*. Fox & Wilkes, New York, 4th revised edn.

VON SCHMOLLER, G. (1883): *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich*chap. Zur Methodologie der Staats- und Socialwissenschaften, pp. 974–994. Leipzig.

WAGEMAN, E. (1943): *Wirtschaftspolitische Strategie: Von den obersten Grundsätzen wirtschaftlicher Staatskunst*. Hanseatische Verlagsanstalt, Hamburg, second rev edn.

WEBER, M. (1949): *The Methodology of Social Sciences*chap. 'Objectivity' in Social Science and Social Policy, pp. 50–112. The Free Press, New York.

Chapter 5

The Transformation Problem: A tale of two interpretations

5.1 Abstract

Over 100 years since Marx's value theory of labour was first published, the "transformation problem" – deriving prices from values and providing a theory of profits as arising from surplus value – continues to present a fascinating challenge to economists of all shades of intellectual persuasion. However, while mainstream economists have by and large come to dismiss the transformation problem as a trivial technical exercise, the issue has recently received renewed attention in heterodox economic theory. This chapter provides a broad historical overview of the transformation problem, highlighting the main ontological similarities and differences of its proposed solutions. Specifically, I discuss (i) how the interpretation of transformation problem has changed over time, (ii) why it was put to rest in mainstream economics and (iii) to what extent it has regained prominence in contemporary Marxian economics.

Keywords: History of economic thought, transformation problem, Marxian economics, value theory of labour
JEL Codes: B1, B2, B41

5.2 Introduction

Over 100 years since the publication of the main elements of Marx's value theory of labour in Volumes I and III in *Das Kapital* (Marx, 1876, 1895), the so-called Transformation Problem – deriving prices from values and providing a theory of profits as arising from surplus value – continues to present a fascinating challenge to economists of all shades of intellectual persuasion. Although it is generally accepted that the transformation problem has been solved (at least mathematically) in a number of ways, there are few problems in economics that have excited similar interest over such a prolonged period. Part of the puzzle's long-standing appeal stems from the fact that transformation problem represents the "most glaring single hole in the Marxian model" (Skousen, 2007, p.93).

The transformation problem may have "been solved and resolved in a number of ways"[1], yet economists' fascination with this topic seems almost strong as today as it was over a century ago. After a several waves of increasing and decreasing attention – the most famous waves perhaps having been triggered by the writings of von Böhm-Bawerk (1898), Winternitz (1948) and then by Samuelson (1971) – there has been a flurry of renewed interest in the topic, primarily spurred by a so-called "new solution" to the transformation problem in recent years.

Indeed, while contemporary mainstream economists have by and large come to dismiss the transformation problem as a trivial technical exercise, the issue has recently received renewed attention in heterodox economic theory. This paper provides a broad historical overview of the transformation problem and specifically focuses on similarities and differences of how the transformation problem has been interpreted, why it was put to rest in mainstream economics and how it has regained prominence in Marxian economics.

The remainder of this paper is structured as follows. Section 5.3 provides an overview of Marx's labour theory of value and section 5.4 describes the transformation problem and illustrates the main solutions that have emerged dealing with this issue. Section 5.5 then puts two contrasting interpretations of these solutions into a historical context. Lastly, section 5.6 offers some concluding remarks.

5.3 Marx's labour theory of value

The labour theory of value (LTV) is a central feature of classical economics and proportionally relates the exchange value of a commodity to the amount of labour that was required to produce it. Although the LTV is perhaps most widely associated with Marxian economics today, its theoretical foundations were laid by Adam Smith and David Ricardo and their ideas were later adopted by Marx without much modification. Indeed, LVT was the predominant theoretical paradigm regarding the determinants of economic value

[1] Desai (1988, p.297).

right until the mid-nineteenth century when it was increasingly replaced by theories of the "marginalist revolution" which linked factor incomes to their marginal products.

5.3.1 Labour as a numéraire

The LTV is a catch-all term that encompasses several theoretical concepts, all of which share the fact that labour serves as a measure of exchange value. Smith's beaver-deer example is perhaps the most well-known illustration of how labour is used as a numéraire. While it would be beyond the scope of this paper to provide a complete review of different varieties of LTV, it is still useful to highlight some of the most important strands of LTV.

Labour demand theory suggest that the exchange value of a commodity is equal to the quantity of labour that can be demanded in exchange. Similarly, *labour cost theory* posits that a commodity's exchange value is determined by the amount of labour that is required in its production. Clearly, if labour is the only factor of production, the labour demand and labour cost approaches are identical.

Recognising that there is no natural numéraire, Ricardo proposes gold as a standard for measuring exchange value and relative prices. While use value is a necessary but not sufficient condition for a commodity's exchange value, he identified scarcity and the quantity of labour embedded in a commodity as the main determinants of exchange value. According to this view, the exchange value of non-reproducible commodities is solely determined by demand, whereas factor inputs, in particular labour, drive the exchange value of reproducible commodities.

Much of Marx's analysis operates within the classical labour cost theory, whereby the exchange value of a good is determined by the amount of labour used in its production. The exchange value of a good is therefore determined by the amount of "socially necessarily labour" used in its production. In contrast to classical theory, however, Marx defines labour as a commodity that is *permanently reproduced*, rather than simply as a factor of production. Because it generally takes less than one unit of labour to produce one unit of labour, a unit of labour produces more than its own subsistence requirement. In other words, the labour cost of labour per unit of labour is less than unity which gives rise to a difference between exchange value of labour and its use value. It is this difference, the so-called *surplus value*, from which all profits arise in the Marxian system. While the value of capital is transferred one-for-one in the production process, only labour is capable of producing surplus value – a Marxian notion of value-added. This forms the heart of Marx's LTV. I will define these terms more formally in section 5.3.3.

In the Marxian system, thus, surplus value only derives from the investment in the employment of labour, not capital. The organic composition of capital however – i.e. the labour-capital ratio – varies across industries, which presents a problem: how then, under competitive conditions, could we reconcile the fact that the rate of profit in different industries would tend to be the same, while the extraction of surplus value depended on the organic composition of capital? *This* is the very essence of the transformation problem. Beyond its apparent contradiction, however, Robbins (1998) sees Marx as reaching much the same conclusion as Ricardo did, "whose theory of value was [...] a real cost theory of value, which admittedly depended upon labour cost, but took account of other factors as well."[2]

In this context, it is important to bear in mind that classical economists – including Marx – did not establish a link between utility and the determination of exchange value.

5.3.2 From values to prices

Beyond the fact that labour is the cause and determinant of value, another important feature of the LTV is the assumption that the own value of labour does not fluctuate. This permits the notion of a natural cost of production which in turn determines the so-called "natural price" of a commodity. While this view does not preclude that the market price of a commodity deviates from its natural price, such deviations are merely viewed as short-run phenomena which do not persist in the long-run. So, contrary to common belief, the LTV does not deny the role of supply and demand influencing prices.

At the same time, the LTV is concerned with the determination of the relative prices of output only and does not explicitly concern itself with the with the remuneration of the individual factor of production. In classical economics, this is the concern of distribution theory which – broadly speaking – looks at how the exchange value can be decompose into factor shares such as wages, rent (in a Ricardian sense) and profits or interest. In contrast to this clear distinction between valuation and distributional aspects of the mode of production, the prices of factors of production and prices of outputs are determined simultaneously within the marginalist framework of the neoclassical paradigm. As a result, value theory and distribution theory are merged. As I will argue later, this is to some extent already anticipated by Marx.

[2] Robbins (1998, p.239).

5.3.3 The Marxian input-output system

In the ensuing section, I loosely follow Samuelson (1983) and Blaug (1997) when outlining the Marxian version of a simplified input-output system of economic production. For the remainder of this paper, I will rely on the following definitions and identities:

c = 'constant capital', depreciation charges on fixed capital and inputs of raw materials,
v = 'variable capital', workers' wages,
s = 'surplus value', excess of receipts over $(c+v)$,
σ = 'rate of surplus value', where $\sigma \equiv s/v$,
r = 'rate of profit', where $r \equiv s/(c+v)$,
q = 'organic composition of capital', where $q \equiv c/(c+v)$.

As mentioned above, the assertion that only labour produces surplus value – which in turn drives profits – lies at the centre of Marx's analysis. Thus it should be possible to determine relative prices from labour values and establish a link between surplus values and profits. As it turns out, however, this is not as straightforward as it sounds as there are some problems with this line of argument. I will discuss these complications next.

Indeed, in Volume I of *Das Kapital*, Marx explained the profits of capital as resulting from surplus value. However, he left open the problem of explaining how capitalists with differing capital-labour ratios, q, can have similar profits. In Volume III, Marx takes up the matter again, and he acknowledges that profit is proportionate to capital rather than labour after all. This "great contradiction" has been at the main focal point for much of the criticism of Marxian economics.

If relative prices are linked to relative labour values and there are no productivity differences, the value of output produced by equal amounts of labour would be the same. With uniform productivity, wage rates are also uniform across industries which implies that the rate of surplus value, σ is the same across all industries. However, different industries clearly have different capital-labour ratios, q, which suggests that profitability, r, should also vary across industries. What is more, a uniform σ and differing qs means that the rate of profit is highest in the most labour-intensive industries – clearly inconsistent with the notion that capitalists substitute capital for labour in search for higher profits.

Furthermore, competition between industries leads to an equalisation of profit rates which – despite the fact that the industry-specific qs are different – cannot be consistent with a uniform σ. This contradiction can be easily visualised as follows:

$$r \equiv \frac{s}{(c+v)} \equiv \frac{\sigma}{(q+1)} \tag{5.3.1}$$

Thus, if the rate of profit, r, is uniform across industries, both σ and q must either be equal or vary inversely. But since q is known to be different across industries, σ must also differ which directly contradicts Marx's own assumption that profits depend only on labour values, i.e. that σ is constant. As relative prices can therefore not correspond to relative labour values, Marx's LTV must be salvaged in some other way, namely by transforming values into prices while maintaining equal rates of surplus and profit across industries. The solution to the great contradiction is the so-called *transformation problem* which is elucidated in more detail in the next section.

5.4 The transformation problem

In light of the previous discussion, Desai (1988) emphasises that the crucial question at the heart of the transformation problem is not so much one of deriving prices from values, but one of providing a theory of profits as arising from surplus values. In the posthumously published Vol. III of *Das Kapital*, Marx himself provided a first solution to the transformation problem, the first step of which is to define the total value produced by department i. Recalling the definitions of variable in the previous section, this can be written as

$$a_i = c_i + v_i + s_i = c_i + (1+\sigma)v_i = \frac{(1+\sigma(1-q_i))}{(1-q_i)}v_i. \tag{5.4.2}$$

The next step is to express the price of output in terms of the industry-wide profit rate as follows

$$p_i = (1+r)(c_i+v_i) = \frac{(1+\sigma(1-q_0))}{(1-q_i)}v_i, \tag{5.4.3}$$

where q_0 is the industry-wide organic composition of capital. A direct comparison of equations (5.4.2) and (5.4.3) gives the basic result of Marx's solution to the transformation problem

$$\frac{p_i}{a_i} = \frac{(1+\sigma(1-q_0))}{(1+\sigma(1-q_i))} \gtreqless 1 \text{ as } q_i \gtreqless q_0. \tag{5.4.4}$$

In other words, the deviation of prices from their corresponding values depends on department i's capital-labour ratio in comparison to the industry average. Using Marx's terminology of referring to different industries as "departments", this implies that if department i is more capital intensive that the average, prices of production will exceed the values of the goods and vice versa.

Table 5.1: The great contradiction

Industry	*Capital*	*c*	*v*	*Cost price*	*s at* $\sigma = 1$	*Value*	*Profit*	*Price of prod.*	*Price > value*
Dept. I	80C + 20V	50	20	70	20	90	20%	92	+2
Dept. II	70C + 30V	51	30	81	30	111	30%	103	-8
Dept. III	60C + 40V	51	40	91	40	131	40%	113	-18
Dept. IV	85C + 15V	40	15	55	15	70	15%	77	+7
Dept. V	95C + 5V	10	5	15	5	20	5%	37	+17
$\sum$	390C + 110V	202	110	312	110	422	22%	422	0

Notes: Following Blaug's notation, capital letters are used for *stocks* and lowercase letters for *flows*. Thus, for example, in the case of Marx's 'constant capital', c is defined as the sum of the depreciation charges on fixed capital and inputs of raw materials, whereas C is the value of the stock of physical equipment, machinery and inventory of raw materials. *Source*: Marx's original example reproduced from Blaug (1997, p.219).

Table 5.1 provides a simple illustration of Marx's solution to the transformation problem by considering an economy with five industries that each have the same amount of capital invested and – by construction – enjoy the same σ. Two elements of this example are particularly worth recalling as the basic result presented in equation (5.4.4). First, for none of the industries does the value correspond to the price of production, although at the aggregate level of the economy the sum of the price-value deviations is zero. Second, the direction of the deviation of the price of production from value depends on the organic composition of capital.

Table 5.2 provides a further illustration of this phenomenon in the context of a simple three-industry case. Here, the organic composition of capital for department II is deliberately constructed in order to correspond to the industry average ($q_2 = q_0$), which is why its value is identical to the price of production. Department I has an organic composition of capital in excess of the industry average, whereas the

Table 5.2: The transformation problem

Industry	*Capital*	*Cost price*	*s at $\sigma = 1$*	*Value*	*Profit $r = 0.33$*	*Price of prod.*	*Price > value*	q_i
Dept. I	250C + 75V	325	75	400	108.3	433.3	+33.3	3.3
Dept. II	100C + 50V	150	50	200	50.0	200.0	0	2.0
Dept. III	50C + 75V	125	75	200	41.6	166.6	-33.3	0.7
$\sum$	400C + 200V	600	200	800	200	800	0	

Notes: The "organic composition of capital", q, is not explicitly defined by Marx, but most commonly assumed to be $\frac{c}{(c+v)}$. However, ratherthan the ratio of two flows, Blaug argues that Marx must have been interested in the ratio of machine cost to labour cost. Therefore, q is defined here as $\frac{C}{v}$. *Source*: Marx's original example reproduced from Blaug (1997, p.220).

opposite is the case for department III ($q_1 > q_0$ and $q_3 < q_0$). Thus with a given σ and industry-wide profit equalisation, for example, department III can only earn the average rate of profit which means that the price of production is less than the value. As is implied in tables 5.1 and 5.2, therefore, the inter-industry equality in the rate of profit causes surplus to be redistributed from labour-intensive industries to capital-intensive industries.

5.4.1 Solving the transformation problem

At first sight, Marx's solution seems to provide a satisfactory theory of profits arising from surplus value. Indeed, prior to the publication of Vol. III, Engels famously issued the challenge to solve the transformation problem, simultaneously announcing that the forthcoming Vol. III would settle all controversy. Yet, he was (partially) wrong as overcoming this paradox still captures economists attention even 125 years after it was first formulated by Marx. The following section provides an overview of some of the main solutions to the transformation problem.

In Marx's description of the transformation problem, industries are not related with each other. However, any general solution of the transformation problem should be able to link all values – both input and outputs – to the corresponding relative prices. In modern terminology, the transformation problem is therefore a *general equilibrium proposition* about how various sectors interact and how the rate of profit is equalised through price of production-value differences as illustrated in the previous section.

5.4.2 Böhm-Bawerk and Bortkiewicz

Perhaps one of the most famous criticisms of the apparent inconsistency in Marx's LTV is due to von Böhm-Bawerk (1898) who maintains that Marx does not resolve the issue logically. The fact that rates of profit rather than surplus value tend towards equality across industries implies that that commodities will sell at their cost of production rather than their labour value. Thus, the process of transforming values into prices necessarily determines the labour theory of value in its entirety.

> "I cannot help myself; I see here no explanation and reconciliation of a contradiction, but the bare contradiction itself. Marx's third volume contradicts the first. The theory of the average rate of profit and of the prices of production cannot be reconciled with the theory of value. This is the impression which must, I believe, be received by every logical thinker ... And even a man who is so close to the Marxian system as Werner Sombart, says that a "general head-shaking" best represents the probable effect produced on most readers by the third volume."[3]

Böhm-Bawerk alleges that Marx's theory of prices of production stands in direct contradiction with the theory that exchange values are determined by the quantity of labour "socially necessary" to produce a commodity. In other words, equal values correspond to equal prices only in industries for which the organic composition of capital is equal to the industry average, i.e. $q_i = q_0$.

Furthermore, going beyond Marx's original analysis of "simple reproduction" where there is no economic interaction between the three departments, the transformation problem also needs to besolved for "expanded reproduction". The mathematical solution to this generalised version of the transformation problem is due to von Bortkiewicz (1907) and was later elaborated by Winternitz (1948).[4] This solution is illustrated below. For this purpose, Department I is now assumed to produce capital goods which are used as intermediate goods by departments II and III. Department II produces wage goods consumed by workers and department III supplies luxury goods for the capitalist class.

Table 5.3 illustrates how this mode of production can be represented by two systems of equations that relate values (c_i, v_i, s_i) to total output (a_i) and prices of production (p_i). Solving the transformation problem can thus be viewed as finding a unique solution relative prices solution.

[3] von Böhm-Bawerk (1898), chapter III, "The Question of the Contradiction".

[4] Historians of economic thought do not uniformly share the same assessment of the relative merits of these solutions. While Desai (1988, p.297), for example, lauds Bortkiewicz's "less turgid and elegant solution", May (1948, p.596) refers to the same efforts as "an artificial confusion [with] pseudo-mathematical mystifications" that bear little relation to the basic problem posed by Marx.

Table 5.3: Another view of the transformation problem

	Values	*Prices of production*
Dept. I	$c_1+v_1+s_1=a_1$	$c_1 p_1+v_1 p_2+\pi_1 p_3=a_1 p_1$
Dept. II	$c_2+v_2+s_2=a_2$	$c_2 p_1+v_2 p_2+\pi_2 p_3=a_2 p_2$
Dept. III	$c_3+v_3+s_3=a_3$	$c_3 p_1+v_3 p_2+\pi_3 p_3=a_3 p_3$
$\sum$	$a_1+a_2+a_3=\sum a$	$cp_1+vp_2+\pi p_3=\sum ap$

Source: Author's adaptation from Winternitz (1948) and Blaug (1997).

Since profits are defined as $\pi_i = r(c_i p_1 + v_i p_2)$ and profitability is assumed to be identical in departments I and II, the rate of profit can be written as $r = \pi_i/(c_i p_1 + v_i p_2)$.[5] This now permits to express the value of total output in terms of the rate of profit as $a_i p_i = (1+r)(c_i p_1 + v_i p_2)$. Rearranging terms, this can then be re-written as

$$1 + r = \frac{a_i p_i}{c_i p_1 + v_i p_2} \text{ for } i = 1, 2. \tag{5.4.5}$$

In his original solution, Winternitz (1948) demonstrates that equation (5.4.5) can be written as the following quadratic equation by re-arranging and setting $m \equiv \frac{p_1}{p_2}$. Thus,

$$m^2(a_1 c_1) + m(a_1 v_2 - a_2 c_1) - (a_2 v_1) = 0, \tag{5.4.6}$$

which has the solution

$$m = \frac{(a_1 v_2 - a_2 c_1) + \sqrt{(a_2 c_1 - a_1 v_2)^2 + (4 a_1 a_2 v_1 c_2)}}{2 a_1 c_2}. \tag{5.4.7}$$

Ignoring negative solutions, m is now given and the average rate of profit is determined as

[5] Department III produces luxury consumption goods and has no influence on average profits by construction.

$$r = \frac{a_1 m}{c_1 m + v_1} - 1. \tag{5.4.8}$$

It is important to note that this solution is somewhat restrictive for two reasons. First, the rate of profit in department III and its organic composition of capital have no influence on the average rate of profit. In other words, this three-industry model assumes that the final use of a commodity is predetermined by the department in which it is manufactured. Second, because the transformation problem has been solved in terms of relative prices (recalling that $m \equiv \frac{p_1}{p_2}$), additional aggregate characteristics are required to determine absolute prices.

Winternitz suggests to choose Marx's proposition that the "sum of prices is equal to the sum of values" (i.e. the labour theory of prices). However, Blaug (1997) notes that there is a second invariance condition that has equal theoretical merit, namely that "the total surplus in value terms is equal to profits in price terms" (i.e. the labour theory of profits).[6]

This now raises the complication that with two invariance conditions and three department equations as displayed in table **??**, the transformation problem is *overdetermined* since there are five equations but only four unknowns (r, p_1, p_2 and p_3). This means that a general solution of the transformation problem is only possible by retaining one of the invariance conditions, but not both. In other words, it is only possible under very restrictive assumptions to solve Marx's original challenge of applying the LVT to both relative prices and profits.

5.4.3 Beyond mathematics

Beyond these mathematical complications outlined in the previous section, the solution of the transformation problem shows that, if value and price are defined according to Marx, there exists a simple transformation connecting the two. The transformation is independent of any equilibrium condition. Thus the transformation problem – in the strictly technical sense of linking value and price of production – is seen to be relatively straight-forward mathematically.

However, even if it can be accepted that the transformation problem can be tackled algebraically, there is a number of other issues that remain unresolved and are keeping economists occupied with this

[6] Blaug (1997, p.225)

issue. A number of these qualitative aspects of solutions to the transformation problem is described in the next section. Many subtle issues regarding the transformation problem reach far beyond its mathematical solvability.

May (1948) points out that the real problem in the context of Marxian economics is not the difficulty of relating values and prices of production, but relating these concepts to *actual prices* which are not expressed in terms of labour time in the production process. This is then one of the weaknesses of the mathematical solution of the transformation problem elaborated by Winternitz (1948) who conflates prices of production with prices at which exchange takes place. Indeed, May highlights that – in Marx's original sense – the price of production is a form of value and that price may not only deviate from price of production, but a commodity may also have a price without having a price of production.

To complicate things further, s and v are not only non-observable, but also non-behavioural variables. This makes it particularly problematic to formulate a priori whether σ, the rate of surplus value, is equal across industries or not. The transformation problem stands traditional economic reasoning and reality on its head in so far as observable prices have to be transformed into unobservable values and not the other way round.

5.5 Two interpretations

By the mid-20th century numerical solutions to the transformation problem were sufficiently generalised and mostly universally accepted. Increasing computing power in the 1960s also made it possible to devise real-world applications to the transformation problem by deriving values and prices of production from actual real prices. At the height of the cold war in the 1960s and 1970s, this was particularly popular practice in COMECON countries. The most ambitious such project took place in Yugoslavia where a group of economists transformed current prices into prices of production on the basis of a 28-sector input-output model (Bajt, 1970, p.371).

Even after these practical applications, there still seemed to remain sufficient substance for debate among economists. Meek (1956) argues that any mathematical or logical solution of the transformation problem would only fill part of the gap in Marx's analysis. To fill the rest of it, one must turn to economic history rather than mathematics. Specifically, he suggested that the "derivation of prices from values" must be regarded as a historical as well as a logical process. As such, the transformation problem can be viewed as a historic description of the evolution of the capitalist laws of motion as opposed to the mode of production in a pre-capitalist society. This aspect of the transformation problem is still actively debated in Marxian economics today and is briefly revisited in section 5.5.2.

In the latter half of the last century, the treatment of the transformation problem was increasingly conducted in two diametrically opposed intellectual camps – mainstream economics on the one hand, and radical, heterodox economists on the other hand. The remainder of this section looks at some of the similarities and differences of how the transformation problem has been treated, why it was put to rest and then re-suscitated, and – most recently – how it has been re-invented.

5.5.1 Transforming the mainstream?

The centenary of the publication of Marx's *Das Kapital* saw a revival of the debate about the transformation problem in mainstream economics. While much of this analysis acknowledges the historical importance of Marx's LTV, there soon was consensus that – with the propositions of marginalist revolution now a mainstay of economic reasoning – the relative prices of goods really do change as demand changes, even when their socially-necessary labour contents do not change. Indeed, in a early reprise on the transformation problem Samuelson (1967) rhetorically concedes that

> "[...] if labour-theory-of-value reasoning, as applied to an impeccable model of equal factor intensities, turned up new light on exploitation in an existing system or if it turned up new light on the laws of development of capitalism, it would be an invaluable tool even though not defensible as a general theory of markets. "[7]

Perhaps not surprisingly, however, he concludes that "Marxian economics is powerless to explain the 1937–1967 developments of European and American economics".[8] Later, Samuelson (1971, 1974) asserts that the transformation problem was in fact a "non-problem" that could be eliminated mathematically. In his view, the transformation problem was then not the Achilles heel of Marxian economics but simply a redundant appendix.[9] Samuelson's focus on the irrelevance of the transformation problem stems from the practical reality that prices and profits could be derived directly from input-output data without the detour of computing values.

At the same time, however, other mainstream economists argued differently and highlighted that the apparent inconsistency between Marx's two theories of value is quite deliberate. While Vol. I might give

[7] Samuelson (1967, p.620).
[8] Samuelson (1967, p.623).
[9] See Bronfenbrenner (1973) for a concise discussion and overview of Samuelson's "vituperative feuding with Marx's ghost".

the impression that the labour values of commodities were intended to explain their exchange values (relative prices), Niehans (1990), for example, asserts that it would be wrong to conclude that Marx was proposing a labour theory of exchange value. Indeed, he highlights that precisely this difference between the two sets of values is one of Marx's major contributions. As such, criticising this deliberate difference between the two (labour account, exchange account) is besides the point.

Thus the question arises whether Marx was simply wrong – or perhaps inconsistent – or why would he deliberately mislead? Baumol (1974a,b) is convinced it is the latter, yet his answer is rather surprising. He maintains that Marx did not intend his transformation analysis to show how prices can be deduced from labour values. Rather, because the two sets of values are derived independently and because they differ in a significant and systematic manner, Marx had set deliberate traps for the express purpose of goading "vulgar economists".

In this view, the relevance of the transformation problem for mainstream economics stems from the general equilibrium insights in which surplus values were redistributed through the economic system via the pricing mechanism to equalise profit rates. Indeed, Baumol argues that this central insight of Marx' LTV had been lost in the increasingly mathematical treatment of the transformation problem, starting with Böhm-Bawerk right up to Samuelson:

> "My contention is that Marx' interest in the transformation analysis as a sequel to his value theory was not a matter of pricing. Rather it sought to describe how *non*-wage incomes are *produced* and then how this aggregate is *redistributed* [...] the substance of Marx' analysis can be summarised in a simple parable, in which the economy is described as an aggregation of industries each of which contributes to a storehouse containing total surplus value [...] if we use labour units to measure these quantities, each industry's contribution is proportionate to the quantity of labour it uses."[10]

5.5.2 The "new solution"

In contemporary Marxist economics, the meaning and significance of the transformation problem goes well beyond a simple mapping of values onto prices and linking surplus value to profits. Broadly speaking, there are two main schools of thought within heterodox economics, each with a very different take on the nature and significance of Marx's LTV and the transformation problem.[11]

[10] Baumol (1974b, p.53), italics in the original.
[11] See Hunt (1989) for a comprehensive overview of this literature.

On the one hand, there is the so-called *rationalist interpretation* that relies on the more philosophical distinction between *essence* and *appearance*. According to this interpretation, the essence of something is never given immediately in its empirical form. Abstract labour, therefore, is the substance whose empirical form is the price. In this approach the issue is not whether one can find a general mathematical solution for transforming abstract labour and value into prices. Indeed, values are only empirically observable as prices.

In contrast to this view, there is the *empirical interpretation* which – accepting that values are the only cause of prices – views the transformation problem as a mathematical specification to account for the divergence of observed prices from values. Thus in its mechanics, though not the assumptions, this school of thought is close to the treatment of the transformation problem within mainstream economics. Out of this paradigm, a new branch of economic literature emerged in the 1980s that deals extensively with more mainstream criticism of the transformation problem of mathematical and logical inconsistency that were discussed in section 5.4.1.

The so-called "new solution" of the transformation problem was first independently put forth by Duménil (1984a,b) and Foley (1984). This new approach rests on treating two aspects differently than in the traditional formulation. First, net value is used instead of gross value to avoid double counting. Second, the division of new value into variable capital and surplus value is determined in terms of money wages paid to workers rather than workers' consumption goods as is the case in the traditional approach.[12] Defining a direct role for money wages in the distribution of surplus value – rather than hypothetical workers' consumption bundles – is perhaps the most significant contribution to the transformation problem by the new solution literature.

Yet, while some hail the new solution as "a very important advance in Marxian scholarship", others are much more cautious and warn that the new solution is premised on theoretical modifications that "cannot be supported easily by textual evidence from Marx's work".[13]

5.6 Outlook

The transformation problem – deriving prices from values and providing a theory of profits as arising from surplus values – and possible solutions to the problem have received widespread attention across

[12] See Campbell (1997, 2002) for a comprehensive review of the literature on the "new solution" to the transformation problem.
[13] See eg. Moseley (2000, p.312) and a more critical assessment of the new solution by Rieu (2006, p.259).

a wide range of theoretical and empirical economics. Unlike only few other issues in the dismal science, it still elicits interest and has the capacity to polarise economists – both in the mainstream and at the fringe. Undeniably, the transformation problem continues to fascinate and antagonise as much today as it did when it was first formulated by Marx over 125 years ago. To many, Eugen von Böhm-Bahwerk's somewhat caustic assessment may still capture this sentiment best:

> "I consider it one of the most striking tributes which could have been paid to Marx as a thinker that this challenge [*the transformation problem*] was taken up by so many persons, and in circles so much wider than the one to which it was chiefly directed … even economists who would probably have been called by Marx "vulgar economists", vied with each other in the attempt to penetrate into the probable nexus of Marx's lines of thought, which were still shrouded in mystery. There grew up between 1885, the year then the second volume of Marx's Capital appeared, and 1894 when the third volume came out, a regular prize essay competition on the "average rate of profit", and its relation to the "law of value". According to the view of Friedrich Engels – now, like Marx, no longer living – as stated in his criticism of these prize essays in the preface to the third volume, no one succeeded in carrying off the prize."[14]

While there may be disagreement among those who claim that they have at least partially succeeded in its solution, there are several aspects of the transformation problem that remain unresolved. As it has been dealt with to date, the transformation problem only presents itself as a static derivation of prices from values and profits from surplus values. It has neither been carried out in dynamic terms yet, nor has it been extended to include monopolistic competition. Furthermore, it has no role for money and does not deal with uncertainty.

Indeed, it can be argued that the transformation problem has been solved by Marx himself and little conceptual progress has been made in the over hundred years since then. While much of the work and effort that has been directed towards the transformation problem has helped to a much clearer formulation and solution of the problem itself, the key aspects of dynamics, uncertainty, and money remain untackled. Indeed a large field for future research that certainly warrants attention from all throughout the dismal science.

[14] von Böhm-Bawerk (1898).

5.A References

BAJT, A. (1970): "A Post Mortem Note on the 'Transformation Problem'," *Soviet Studies*, 21(3), 371–374.

BAUMOL, W. J. (1974a): "The Fundamental Marxian Theorem: A Reply to Samuelson: Comment," *Journal of Economic Literature*, 12(1), 74–75.

——— (1974b): "The Transformation of Values: What Marx "Really" Meant (An Interpretation)," *Journal of Economic Literature*, 12(1), 51–62.

BLAUG, M. (1997): *Economic Theory in Retrospect*. Cambridge University Press, Cambridge, U.K., fifth edn.

BRONFENBRENNER, M. (1973): "Samuelson, Marx, and Their Latest Critics," *Journal of Economic Literature*, 11(1), 58–63.

CAMPBELL, A. (1997): "The Transformation Problem: A Simple Presentation of the "New Solution"," *Review of Radical Political Economics*, 29(3), 59–69.

——— (2002): "The Nature of Surplus Value in the "New Solution"," *Review of Radical Political Economics*, 34(1), 69–72.

DESAI, M. (1988): "The Transformation Problem," *Journal of Economic Surveys*, 2(4), 295–333.

DUMÉNIL, G. (1984a): "Beyond the Transformation Riddle: A Labor Theory of Value," *Science & Society*, 47, 427–450.

——— (1984b): "The So-called "Transformation Problem" Revisited: A Brief Comment," *Journal of Economic Theory*, 33, 340–348.

FOLEY, D. K. (1984): "The Value of Money, the Value of Labour Power and the Marxian Transformation Problem," *Review of Radical Political Economics*, 14(2), 37–47.

HEILBRONER, R. L. (1999): *The Worldly Philosophers: The Lives, Times And Ideas Of The Great Economic Thinkers*. Touchstone, seventh edn.

HEILBRONER, R. L., AND W. S. MILBERG (1995): *The Crisis of Vision in Modern Economic Thought*. Cambridge University Press, Cambridge, UK, first edn.

HUNT, E. K. (1989): "The Meaning And Significance Of The Transformation Problem: Two Contrasting Approaches," *Atlantic Economic Journal*, 17(4), 47–54.

MARX, K. (1876): *Das Kapital: Kritik der politischen Ökonomie*, vol. I – Der Produktionsprozeß des Kapitals. Dietz Verlag, Berlin, 2005 edn.

——— (1895): *Das Kapital: Kritik der politischen Ökonomie*, vol. III – Der Gesamtprozeß der kapitalistischen Produktion. Dietz Verlag, Berlin, 2005 edn.

MAY, K. (1948): "Value and Price of Production: A Note on Winternitz' Solution," *Economic Journal*, 58(232), 596–599.

MEEK, R. L. (1956): "Some Notes on the "Transformation Problem"," *Economic Journal*, 66(261), 94–107.

MOSELEY, F. (2000): "The "New Solution" to the Transformation Problem: A Sympatheic Critique," *Review of Radical Political Economics*, 32(2), 282–316.

NIEHANS, J. (1990): *A History of Economic Theory: Classic Contributions 1720–1980*. Johns Hopkins University Press, Baltimore and London.

RIEU, D.-M. (2006): "A Reexamination of the Quantitative Issues in the New Interpretation," *Review of Radical Political Economics*, 38(2), 258–271.

RIMA, I. H. (2001): *Development of Economic Analysis*. Routledge, London, sixth edn.

ROBBINS, L. (1998): *A History of Economic Thought: The LSE Lectures*. Princeton University Press, Princeton and Oxford.

SAMUELSON, P. A. (1967): "Das Kapital: A Centenary Appreciation – Marxian Economics as Economics," *American Economic Review*, 57(2), 616–623.

——— (1971): "Understanding the Marxian Notion of Exploitation: A Summary of the So-Called Transformation Problem Between Marxian Values and Competitive Prices," *Journal of Economic Literature*, 9(2), 399–431.

——— (1974): "Insight and Detour in the Theory of Exploitation: A Reply to Baumol," *Journal of Economic Literature*, 12(1), 62–70.

——— (1983): *Foundations of Economic Analysis*, Harvard Economic Studies. Harvard University Press, Cambridge, MA, enlarged edn.

SKOUSEN, M. (2007): *The Big Three in Economics: Adam Smith, Karl Marx and John Maynard Keynes*. M. E. Sharpe, Armonk, NY.

VON BÖHM-BAWERK, E. (1898): *Karl Marx and the Close of His System*. T.F. Unwin, London.

VON BORTKIEWICZ, L. J. (1907): "Wertrechnung und Preisrechnung im Marxschen System," *Archiv für Sozialwissenschaft und Sozialpolitik*, 34(3), 370–385.

WINTERNITZ, J. (1948): "Values and Prices: A Solution of the So-called Transformation Problem," *Economic Journal*, 58(230), 276–280.

Part III

Future Work

Chapter 6

Spatial characteristics of urban specialisation and concentration in the United States☆

6.1 Abstract

In this chapter, I am investigating an alternative regional classification of regions for MSAs on the basis of clusters that were formed by principal component analysis from economic variables that are relevant for regional growth. These variables include labour productivity growth, measures of local industry mix, human capital, entrepreneurship and innovation. I then use these growth-based regional clusters to control for the presence of cluster-specific fixed-effects when explaining the spatial characteristics of urban specialisation and concentration in the United States. In particular, this chapter explores the advantages and disadvantages of urban specialisation and diversity in the context of the US high-tech industry.

Keywords: Spatial patterns of growth, metropolitan growth, specialisation, diversification
JEL Codes: R12, R39, C25

☆Several parts of this work were jointly developed with Heike Mayer (Dept. of Economic Geography, University of Bern).

6.2 Introduction

Unlocking the black box that contains the mechanisms responsible for economic growth and regional development has long been a coveted prize in theoretical and applied regional science. Even the partial identification of the determinants of growth would improve policy design, promising greater potential for improving local economic conditions and attenuating regional business cycle fluctuations. While the impetus for examining the origins of regional growth is clear, firm agreement on what constitutes a region remains elusive. Since the 1950s the Bureau of Economic Analysis (BEA) has grouped the states into eight regions based primarily on cross-sectional similarities in their socioeconomic characteristics. Recognizing the limitations of this regional classification scheme, several recent studies have looked to further the understanding of regional composition.

Crone (2005) and Crone and Clayton-Matthews (2005) group states into regions based on the similarities in their business cycles. They apply k-means cluster analysis to the cyclical components of Stock-Watson-type indices estimated at the state level to group the 48 contiguous states into eight regions with similar cycles. Most recently, O'hUallacháin (2008) uses principal components and cluster analyses as a framework for the identification of regions based on state-level growth measures in the US. Similarly, Owyang, Rapach, and Wall (2009) model the US business cycle using a dynamic factor model that identifies common factors underlying fluctuations in state-level income and employment growth.

At the same time, it has long been recognised that the sectoral composition of urban economies is a key determinant of their economic performance. Chinitz (1961) famously observed that larger cities are more diversified than smaller ones, putting "the whys and wherefores of urban diversification" at the center of inquiry right at the outset of the rapid regional transformation of the post-war US economy. Despite ample evidence of how industrial bases vary across cities and how they vary by city size, our understanding to what extent the structure of cities and the activities its firms and households change over time remains limited. How does the sectoral composition of cities in influence their evolution? Henderson's (1974) canonical model of urban size has recently been extended by Duranton and Puga (2001) and Duranton (2007) to provide the microfoundations of sectoral urban diversity. Although this work offers new insights by showing how innovation shapes the corresponding growth dynamics of urban rise and decline more formally, detailed empirical evidence of how the growth of one industry in an area affects the suitability of as a location for another industry is still sparse.

The aim of this chapter is to outline an agenda for future work along two interrelated strands. The first aspect of this work attempts to further narrow the definition of a region by using indicators of economic growth, population, and social capital at the level of the metropolitan statistical area (MSA), specifically taking advantage of the BEA's recently-released MSA-level GDP data (Panek, Baumgardner, and McCormick, 2007). In the second aspect of this work, I plan to explore the relationship between differences in specialisation and sectoral concentration and contrasts in agglomeration among US metropolitan areas. Looking specifically at the high-tech industry, I will then investigate what size cities high-tech industries tend to gravitate towards, examining the explanatory power of these new regional clusters in this process.

6.2.1 The geography of specialisation and concentration

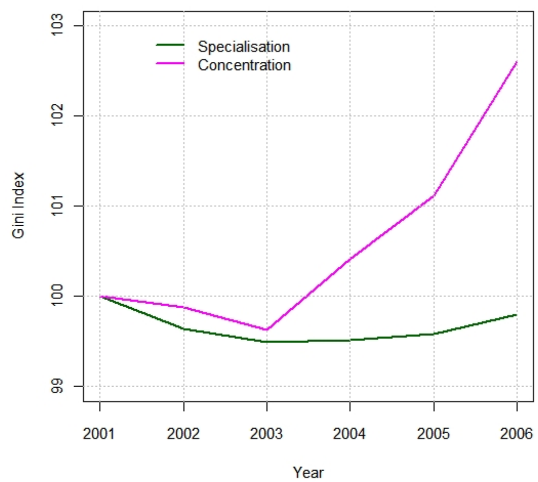
Specialization of regions in particular sectors and concentration of industries in regions or countries have a long tradition of being treated as closely related economic phenomena. In many instances, these two closely-related economic phenomena are used interchangeably when analysing the localisation of economic activities. Most recently, Cutrini (2009) proposes an entropy-based measure of overall localisation that allows specialization to be conceptualized as the mirror image of concentration. However, there a considerable body of evidence that suggests that the specialisation of regions has tended to follow different trends compared to the regional concentration of industries.[1] Indeed, Rossi-Hansberg (2005) proposes a model which predicts that specialization and concentration do go in opposite directions when transport cost change. Using data from the US and EU, Aiginger and Rossi-Hansberg (2006) show that specialization and concentration do not develop in parallel by replicating some of the predicted features of specialisation-concentration divergence predicted.

In the context of the high-tech industry in the United States, the spatial characteristics of specialisation and concentration were illustrated in figures 2.1 and 2.2 of chapter 2. Recalling that metropolitan high-tech concentration (as measured by the location quotient) and its specialisation counterpart (as proxied by the percentage of urban GDP from high-tech) are closely related, however, figure 6.1 highlights that the two measures of localisation have moved in opposite directions from 2001 to 2006. In order to calculate specialization and concentration respectively, I use GDP data published by the Bureau of Economic Analysis for the 358 MSAs of the contiguous US and for ten NAICS-based industries.

In line with the theoretical predictions, I find evidence that the average specialisation among US

[1] See Aiginger and Davies (2002) for a comprehensive overview of this literature.

Figure 6.1: Gini indices of specialisation and concentration for metropolitan GDP, 2001–06



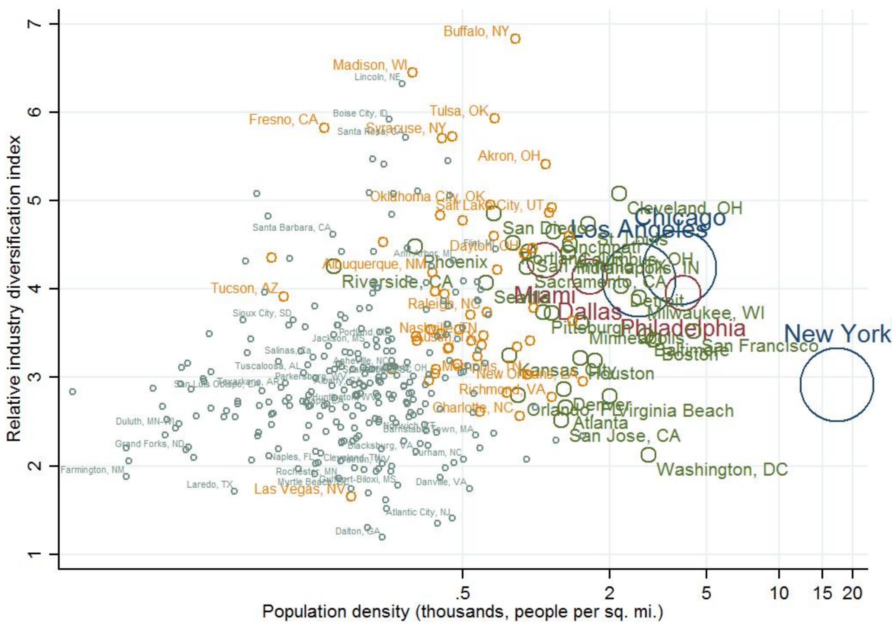
Source: Author's calculations using a NAICS-based definition of high-tech industries following Hecker (2005) and Bureau of Economic Analysis (BEA) for the metropolitan GDP data (Panek, Baumgardner, and McCormick, 2007).

metropolitan areas marginally decreased – as measured by the Gini index – from 0.747 in 2001 to 0.745 in 2006, or by -0.2%. Over the same period, the Gini coefficient for concentration of industries in the metropolitan portion of the US on the other hand increased from 0.448 to 0.460 or by 2.6%. Thus specialisation decreases and concentration increases as predicted by the Rossi-Hansberg model.

While a variety of mechanisms could account for these observed changes in agglomeration, different theories make different predictions about which industries should be coagglomerated in which regions. Chapter 2 showed that one can broadly distinguish between three types of agglomeration economies in the context of knowledge-based innovation: (i) localised learning (Glaeser, 1999), (ii) product and process development cycles (Saxenian, 1994), and (iii) quality-based competition (Porter, 2000). This chapter does not attempt to distinguish between different types of agglomeration effects, but assume that spillovers can be due to any one of these forces. However, localised learning is the main channel through which agglomeration economies is expected to operate in the present context.

While the average diversification of US cities has decreased, larger cities still tend to be more diversied. Cities of similar specialisation are of similar size as it is still widely believed that density-specific agglomeration externalities are among the primary determinants of this diversification (e.g. Duranton and

Figure 6.2: Industry diversification and population density of US metropolitan areas



Note: The relative-diversity index (RDI) is computed by summing for each city i, over all j sectors, the absolute value of the difference between each sector's share s_{ij} in local employment and its share in national employment s_j. Formally, this yields $RDI_i = 1/|\sum_j s_{ij} - s_j|$ (cf. Duranton and Puga, 2000). Population density in thousands of inhabitants per square mile is plotted on a logarithmic scale.

Puga, 2000). Figure 6.2 illustrates this positive correlation between city density and a relative-diversity index, calculated on the basis of all private employment at the two-digit NAICS level in 2007. This link is not especially strong, however, partly because the population density in most larger cities is not exclusively driven by productivity differentials, but also – as we have seen in chapter 3 – by non-tradable amenity differences that both influence the quality-of-life and the quality of the business environment.

The remainder of this chapter is structured as follows. Section 6.3 describes the data set and performs some exploratory spatial analysis. Section 6.4 reports the results from constructing the regional indices and section 6.5 applies these indices within the discrete choice framework for firm location choice that was presented in chapter 2. Section 6.6 then looks at a variant of the previous framework which is used to study urban specialisation and concentration in the high-tech industry, briefly reporting some preliminary empirical results. This is followed by a short overview of future research on this topic in section 6.7.

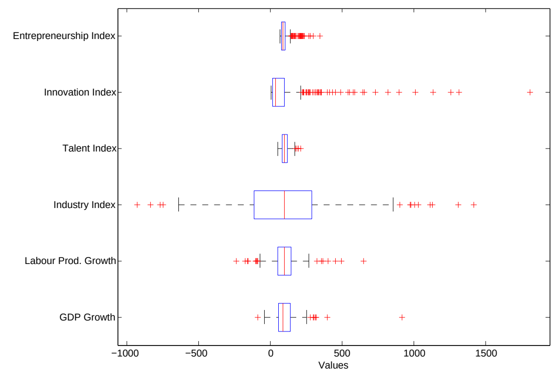
6.3 Data and calibration

The quest to determine the drivers of regional growth is a old as regional science itself. Increased economic linkages between regions is shaping the economic fortunes of both of those at the core and the those at the periphery. This raises a series of questions about the spatial pattern of regional activity. Why are economic activity and prosperity spread so unevenly across space? Do trade and spatial interaction necessarily narrow these differences? What explains the discrepancy between the predictions of theory and what happens in reality? There are a variety of well-established theoretical frameworks that deal with these analysis related to these questions (cf. also chapter 1 on regional linkages). In this context, a great number of factors have been put forward as key determinants of regional economic activity: agglomeration economies, transportation, transaction and other input costs, changes over the product or the spatial life cycle or sources of locational competitive advantages to name just a few of the most important ones.

6.3.1 Key categories of economic activity

These different theories provide the framework for constructing a data set of socio-economic variables that are available at the MSA-level. In a first step, I group the raw data of individual variables into five thematic classes and then construct equally-weighted indices for each group which form the basis for my analysis. These categories of economic activity and the indices are listed below:

Figure 6.3: Box plot of normalised data set



- *Economic Growth*: MSA-level GDP-growth and labour productivity growth (2001-2005)
- *Industry*: Index of employment, establishments, industry concentration, dispersion (income vs. value-added)
- *Talent*: Index of "creative class" and other measures of human capital
- *Entrepreneurship*: Index of firm formation, share of small firms, self-employed labour force, SBIR grants
- *Innovation*: Index of R&D funding, R&D employment and patents

Gross domestic product (GDP), population, innovation, entrepreneurship and creative class measures across the 360 MSAs from the 48 contiguous states were compiled from the Bureau of Economic Analysis, the Bureau of Labor Statistics, the Census Bureau, and the USDA Economic Research Service. MSAs are measured as county-based areas with at least one urbanized area with a population of 50,000 or more, plus adjacent territory that has a high degree of social and economic integration with the core, as measured by commuting ties. Figure 6.3 shows the large variation of the data set across the different measures.

6.3.2 Empirical strategy

My empirical strategy can be broadly described in four steps. First, I perform an exploratory investigation of spatial properties of the data. Then, the dimensions of the data set are reduced by means of principal component analysis (PCA) in order to derive independent key drivers. Similarly, I also perform a factor analysis to check for the stability of the results and to see if any additional insights can be gained from factor scores with maximum explanatory power instead of just obtaining minimum least-squares transformations. The principal components and factors then form the basis for the cluster analysis with the aim to create different regional groups of MSAs. Lastly, I am bringing my data to the model by examining the relationship between the growth-based regional cluster indicator variables with clusters that are formed on the basis of industrial specialisation and sectoral concentration.

The substantial regional variation in the distribution of the basic economic indicators is also illustrated in figure 6.4. Here, the height of the extrusions indicates the level of the index and the colour shading indicates MSA-level annualised compound labour productivity growth which ranges between -5.5% (dark red) and 15.3% (dark green) between 2001 and 2005.

6.3.3 Spatial analysis

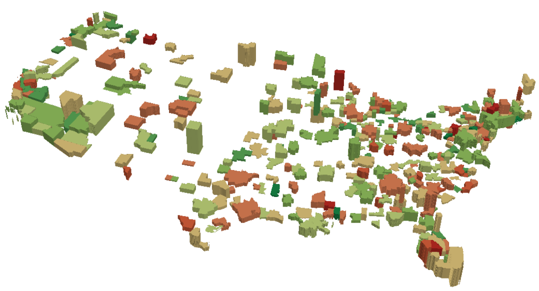
As a first step when investigating the spatial patterns and dependencies of the data set, the correlogram in figure 6.5 reveals that – while there is only mild correlation amongst most variables – talent, innovation and high-tech firm births are significantly positively correlated. It is precisely this co-dependency that I am investigating later when relating firm location choices to these variables in the discrete model framework. In addition to the correlogram, I also constructed a directional semivariogram which did not reveal any significant spatial dependencies for any of the variables.

Similarly, the presence of statistically significant spatial autocorrelation amongst the variables can also be excluded on the basis of a Moran's I test. The results for this test are reported in table 6.1.

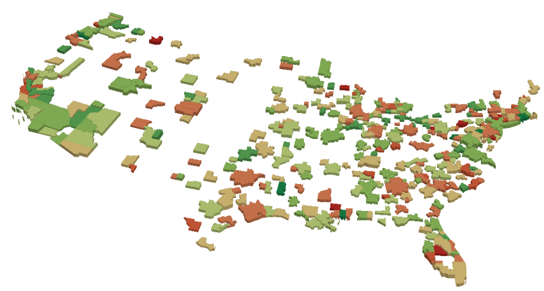
6.4 Constructing regional economic indicators

The regional economic indicators are constructed using different methods of factor analysis in order to extract a set a factors from the data set. In a first step, I am performing a principal component analysis which transforms the data such as to derive independent, uncorrelated observations. In a next step and

Figure 6.4: Spatial distribution of aggregate indices

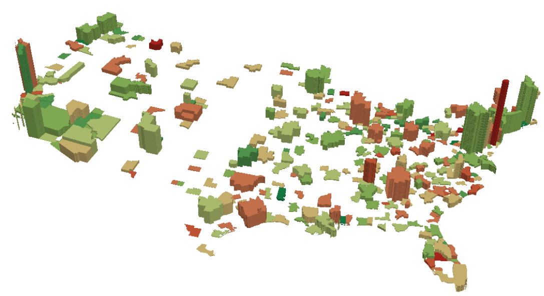


(a) Industry



(b) Talent

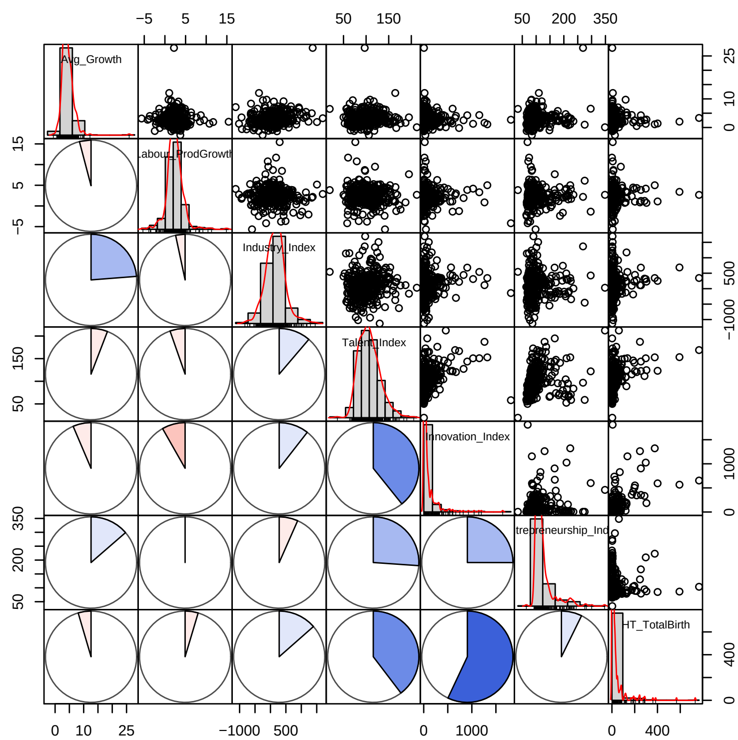
(c) Entrepreneurship



(d) Innovation

Notes: The indices were formed on the basis of individual variables from which five equally-weighted indices were constructed for each group. The height of the extrusions indicates the level of the index and the colour shading indicates MSA-level annualised compound labour productivity growth ranging -5.5% (dark red) and 15.3% (dark green) between 2001 and 2005.

Figure 6.5: Correlogram for key variables



Notes: The variable in the first column is average MSA real GDP growth from 2000 to 2006, the second column is metropolitan labour productivity growth, the third variable is the industry structure index, followed by indices that capture talent, innovation and entrepreneurship. The variable in the last column is the total number of high-tech firm formations between 1998 and 2006. The size of individual slices indicates the strength of bivariate correlation between any two variables, where positive correlations are in blue and negative correlations are shown in red.

Table 6.1: Moran's I test under randomisation

Variable	*Moran's I*	*Expectation*	*Variance*
GDP GROWTH	0.0029	-0.0028	0.0004
PROD. GROWTH	0.0440	-0.0027	0.0004
HT BIRTHS	-0.0118	-0.0028	0.0003

Sources: Author's calculations in R using the `spdep()` package.

as a control for the stability of the PCA results, I conduct a factor analysis, where the factor rotation is chosen such that a small set of factors with large loadings explains most of the variability among the economic random variables in data. Different types of cluster analysis are then performed on the basis of the factor scores from the PCA and the factor analysis to from groups of MSAs.

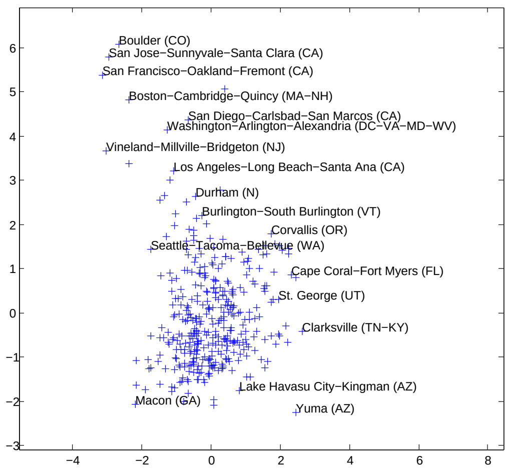
6.4.1 Principal component and factor analysis

Panel (a) in figure 6.6 plots the distribution of the MSAs against the first two principal components. While the majority of MSA seem to form a relatively dense cloud of data points, the group of the largest MSAs is found on the fringes of the main data mass. This is hardly surprising since, for example, the most populous MSA (New York-Northern New Jersey-Long Island) is an economic powerhouse that accounts for almost 10% of GDP among all MSAs. The biggest ten MSAs account for an impressive 38% of total MSA GDP.

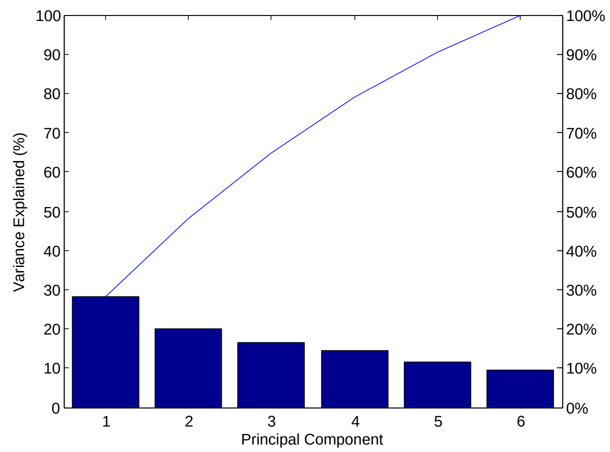
As can be seen from panel 6.6(b), the first three principal components explain almost two-thirds (64%) of the total variance in the data, while all of the variance is explained by six components. A biplot of the factor scores for the main variables and the factor loadings for the first three components is shown in figure 6.7(a).

As an alternative to constructing orthogonal components, the economic indices could be decomposed in to factors that explain the maximum variance of the data set. To obtains such factors, the data is essentially by rotated in such a fashion as to optimise a given variance objective function. I use the *varimax* method since it ensures that that each factor has a small number of large loadings and a large number of zero (or small) loadings. This generally simplifies the interpretation because, after a varimax rotation, each original variable tends to be associated with one (or a small number) of factors, and each factor represents only a small number of variables.

Figure 6.6: Principal component analysis

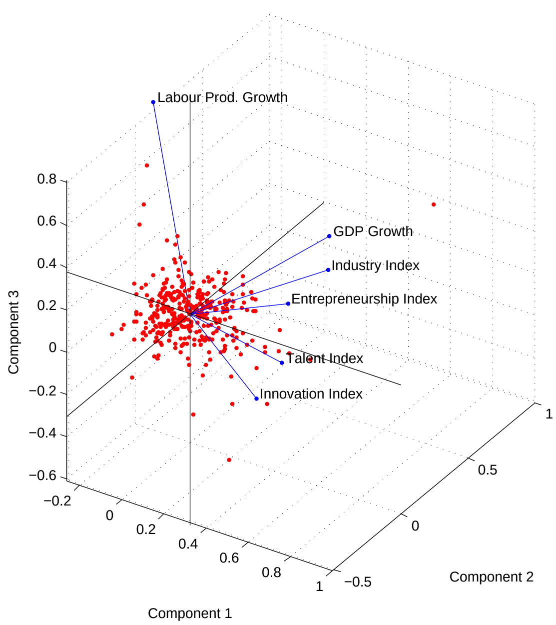


(a) First PC projection of combined data set

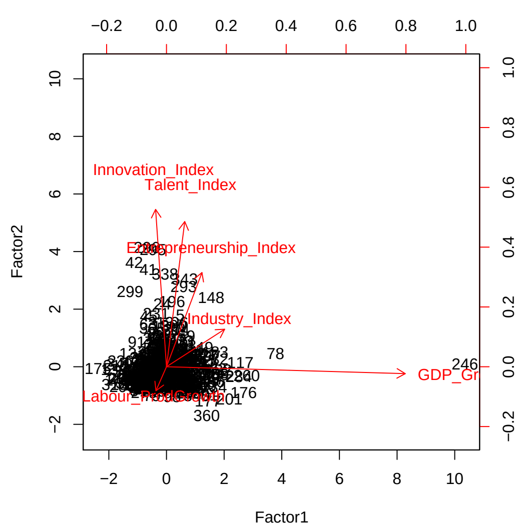


(b) Pareto plot of variance explained

Figure 6.7: Factor analysis and index scores



(a) PCA biplot of MSA index component scores



(b) Factors, scores and factor loadings

In order to make this part of the analysis directly comparable to the PCA, a restriction for same number of factor (six) was imposed. The results of these are visualised in the panel on the right-hand side of figure 6.7(b). A first inspection reveals that both the PCA and factor analysis yield similar results.

6.4.2 Cluster analysis: Numerical taxonomy of regional groups

Hierarchical cluster analysis was used to group MSAs on the basis of the factors identified in the PCA and factor analysis. Observations are grouped together to form sets that are basically homogenous and distinct from other sets. This method plays an important role in the classification of statistical data and is a useful first step in the development of theories that seek to explain the composition of regions based on economic activities.[2]

A hierarchical method is appropriate when dealing with spatial data, since non-hierarchical methods exclude distance measures which are thus unsuitable for identifying regions whose formation is likely influenced, at least in part, by relative proximity. Use of either method requires a clear acknowledgement of inherent subjectivity and distortion due to quasi-arbitrary parameter selection. To be consistent with the analysis of Crone (2005), I am also performing cluster analyses using non-hierarchical methods (k-means, fuzzy clustering) which yield similar results.

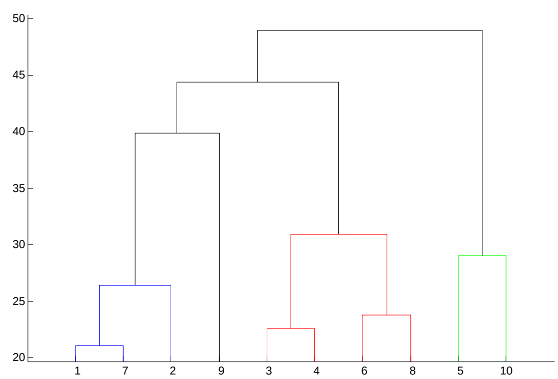


Figure 6.8: Dendrogram from hierarchical clustering (complete linkages)

The sorting strategy and distance method potentially significantly influence outcomes. In an attempt to minimize associated incongruities, I chose the Mahalanobis distance for the construction of the dis-

[2]See O'Sullivan and Unwin (2003) for an extensive treatment of this point.

tance matrix and then used complete linkage (furthest neighbour) as the sorting strategy. Figure 6.8 displays the dendrogram of clusters that was obtained. In order to obtain results that are easily compared with the BEA regions and in line with previous work, I am constraining the number of clusters to eight. A further rationale for this choice is based on the partial hypothesis that MSA growth rates might be homogenous within the 48 states or within the 8 regional groups defined by the Bureau of Economic Analysis. Indeed, the respective dendrograms provide some evidence that could be consistent with this assumption.

Given the varying scales within the data used and the likelihood of some level of correlation amongst variables the Mahalanobis distance method was determined to be most suitable. The Mahalanobis distance varies from both the Euclidean and Manhattan distances in that it accounts for correlation within the data and is scale-invariant. Use of either Euclidean or Manhattan distances would like result in a bias toward the largest values present in the data set. Economic regions may not be geographically contiguous and may consist of a number of MSAs thereby ruling out the use of a single linkage strategy. Distance matrices are known to cluster more successfully than correlation matrices as they yield higher cophenetic correlations and are less susceptible to dramatic changes among different clustering methods which further strengthens the case for using the complete linkage strategy.

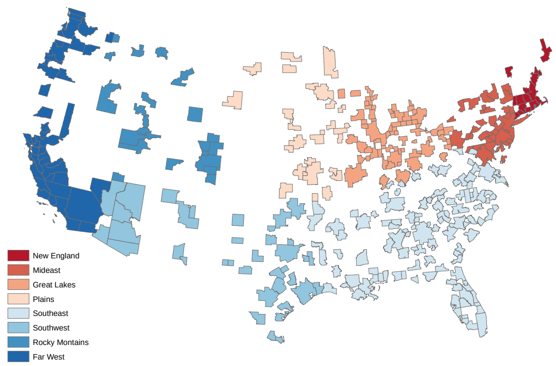
The cophenetic correlation (CC) for a cluster is defined as the linear correlation coefficient between the cophenetic distances obtained from the dendrogram, and the original distances (or dissimilarities) used to construct the tree. Hierarchical and fuzzy clustering yield the highest CCs with 0.8227 and 0.8512 respectively. This means that there is only a minimal degree of distortion due to the sorting strategy which is reflected in the equal distribution of heights at which the links occur in the hierarchy of the dendrogram.

6.4.3 Clusters across BEA Regions

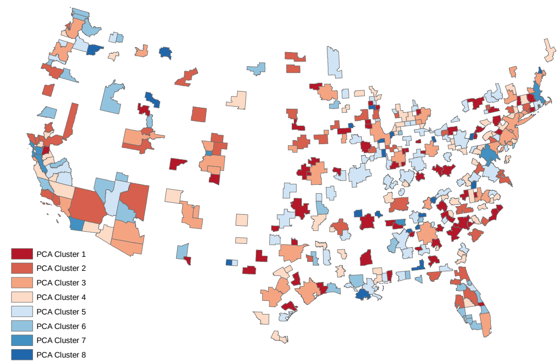
It is not surprising that the BEA's regional classifications do not appear to coincide with the empirical results of the clusters obtained here. While the BEA definitions strongly rely on geographical criteria when defining a region, I use almost exclusively economic data as inputs for the classification. These differences are shown in figure 6.9 where the left panel displays the BEA regions and the right panes plots the regions as defined on the basis of the PCA scores.

Nonetheless, the mosaic plots in figure 6.10 confirms that the spatial distribution of growth and creativity display several interesting (dis)similarities in cluster behaviour across the standardised BEA re-

Figure 6.9: Spatial distribution of regional clusters



(a) MSAs by standardised BEA regions



(b) Clusters across BEA regions

gions. Furthermore, my analysis highlights other interesting avenues for further inquiry, in particular the fact that additional insights could be gained from exploring the temporal behaviour of these clusters.

In the next section, I briefly illustrate the application of these growth-based regional clusters by using them inputs into the model of firm location choice in section 2.5 of chapter 2. The main hypothesis here is that the location of new firms varies regionally. Thus controlling explicitly for regional unobserved characteristics in addition to existing control variables might improve the overall fit of the model.

6.5 Application: High-tech firm location choice

Table 6.2 presents the estimates for the location determinants of high-tech firm births, using both the DM and CLM regressions. While the likelihood ratio test provides evidence for overdispersion, the parameter estimates vary only slightly between both models. With the exception of the control for the cost of labour, all estimates are significant across all estimations. Columns (1) to (2) present the CLM results, including the estimation with the creative class measure and the regional cluster dummy variables. Columns (3) and (5) show the results using the DM which accounts for overdispersion. Columns (1) and (3) corresond to columns (3) and (5) of table 2.7 in chapter 2. Adding the cluster dummies improves the overall fit of the model and yields more similarity of the parameters estimates between both models. The majority of the cluster dummies are significant (not reported here) which provides strong evidenced for the presence of *cluster-specific fixed-effects* in the context of firm's location choices.

Overall, I am able to find supporting evidence that both location factors and cost factors have a signif-

Table 6.2: Determinants of high-tech plant openings with regional clusters

	Model			
	CLM	CLM	DM	DM
Variables	(1)	(2)	(3)	(4)
LABOUR COST	0.0399*	-0.3337	-0.3033	-0.4226
	(0.077)	(0.081)	(0.101)	(0.113)
LAND COST	-0.4981	-0.4558	-0.5319	-0.4225
	(0.047)	(0.054)	(0.060)	(0.065)
TAXES	-0.1122	-0.1315	-0.0717	-0.0927
	(0.023)	(0.024)	(0.028)	(0.028)
WEIGHTED MARKET SIZE	1.0779	1.0988	1.0743	1.0809
	(0.013)	(0.014)	(0.016)	(0.016)
LOCALISATION ECONOMIES	0.0215	0.0227	0.0185	0.0189
	(0.001)	(0.001)	(0.001)	(0.002)
URBANISATION ECONOMIES	0.2827	0.2492	0.3330	0.2346
	(0.045)	(0.044)	(0.055)	(0.044)
HUMAN CAPITAL†	0.0028	0.0016*	-0.0121	0.0006*
	(0.002)	(0.002)	(0.004)	(0.003)
NATURAL AMENITIES	0.2096	0.1334	0.1071	0.1269
	(0.032)	(0.041)	(0.040)	(0.039)
WEIGHTED UNIVERSITY R&D	0.0036	0.0027	0.0029	0.0021
	(0.001)	(0.001)	(0.001)	(0.001)
CREATIVE CLASS	0.5335	0.9618	0.4071	0.2354
	(0.097)	(0.297)	(0.115)	(0.096)
REGIONAL CLUSTERS	–	Yes	–	Yes
Log-likelihood	-11,525.9	11,549.7	-10,929.3	-10,921.3
LR overdispersion test χ^2			994.4	1,093.6
Radius in miles (δ)			60	60
N.Obs.	21,469			

Notes: Estimated models are the conditional logit model (CLM) and the Dirichlet-Multinomial model (DM). Standard errors in parentheses. † Human capital is proxied by the share of university degree holders in the specification of columns (2)–(4). All estimates are significant at the 1% level, except those marked by *. *Sources*: Author's calculations.

Figure 6.10: Mosaic plot of clusters accross BEA regions

(a) Clusters from factor analysis

(b) Clusters from principal components

icant impact on the number of new high-tech firm openings. The estimated coefficient for the spatially weighted market size also proves to be statistically significant and has the expected sign. There is also evidence that the two agglomeration variables – LOCALISATION and URBANISATION – affect high-tech investors locational decisions. Furthermore, the presence of qualified labour and natural amenities also have an influence on new firm formation. One of the original key variable of interest, UNIVERISTY R&D, is positive and statistically significant, suggesting that a closely situated university plays a important role in explaining high-tech location decisions.

In contrast to the overall results, not all “milieu variables” enter significantly across industries. While URBANISATION ECONOMIES and QUALIFIED show an alternating pattern of significance in the different industries, CREATIVE CLASS is significant in all but the chemical industry (SIC 28). Apart from industry-specific differences, one possible alternative explanation is certainly the high levels of correlation among those there regressors. This is something that would need further inquiry. Indeed, this suggests that in addition to using separate measures for urbanisation and the level of human-capital, further insights can be gained in to the firm location decision process by adding a measure that proxies for creativity. Initial explorations for the industry-level data indicate similar results from the DM model, thus suggesting that overdispersion may not be the key driver of these results.

The following section looks at a variant of the previous framework. While largely intended as roadmap for future work, it outlines a number of elements in the context of urban specialisation and concentration in the high-tech industry across metropolitan areas in the United States.

6.6 Specialisation and concentration in the high-tech industry

Technology-based economic development and growth is spreading beyond Silicon Valley and Boston's Route 128. These two pioneering high-technology regions have long captured the attention of policy makers and analysts and many wondered what it would take to become "the next Silicon Valley". So far, efforts to imitate Silicon Valley have had a dismal track record. Now, however, other regions are gaining momentum and are emerging as high-technology locations, utilizing their own ways to enter the knowledge economy. Places like Portland, Boise and Kansas City host significant concentrations of high-technology industry activity.[3] Relative to their size and location, these regions are highly innovative and entrepreneurial. Building on corporate assets, these emerging high-technology regions have recognized the potential of knowledge-based economic growth and are developing unique policies to link universities with industry, connect to other regions in the nation and abroad, facilitate entrepreneurship, and support the development and commercialization of innovation. This section measures recent trends in high-technology growth and develops a typology of high-tech regions.

6.6.1 Technology and industry structure

As our regional economies change into more knowledge-based, innovation-driven and service-oriented economies, it is important to explore the extent to which high-technology industries develop beyond Silicon Valley and Boston. This study examines second tier regions and the characteristics that led to their successful emergence as high-technology locations. To do so, we employ a comprehensive assessment of the 358 metropolitan areas in the contiguous United States along various dimensions of knowledge-based economic growth, including high-technology employment, talent, innovation, and entrepreneurship.

[3] The history of high-tech development in Portland, Boise City and Kansas City, their potential as future leading hubs for the high-tech industry as well as policy options for economic development are discussed in Mayer (2009).

There are two very different approaches to studying market structure. One looks to "industry characteristics" to explain why different industries develop in different ways; the other looks to the pattern of firm growth within a "typical" industry to describe the evolution of the size distribution of firms. Sutton (2001) sets out a unified theory that encompasses both approaches, while generating a series of novel predictions as to how markets evolve. See Scherer (2000) for a critical review. Lehto (2007) for evidence of regional impacts such a geographic and technological proximity on total factor productivity.

6.6.2 Defining high-technology industries

Defining high-technology industries is not an easy task. Analysts typically distinguish between output and input-based definitions. Output-based definitions, such as the one used by the American Electronics Association (AeA), consider an industry high-tech if its products embody technologies. AeA, for example, defines an industry as high tech if it is a "maker/creator of technology, whether it be in the form of products, communications, or services" (AeA, 2003). The selection typically involves industry experts and can be quite labor intensive and subjective. A more common and less subjective approach is to define high-technology industries by the inputs needed to create innovative and technology-oriented products and processes. The two most common measures are employment of engineers, scientists and technicians and investments in research and development. Typically a threshold is defined at which an industry becomes high tech.

For this study, I utilize an employment-based definition of high-technology industries and follow the approach of Hecker (2005). Hecker selected eight technology-oriented occupations including engineers, technicians, life and physical scientists, engineering and natural science managers. He then used the Bureau of Labor Statistics' National Employment Matrix to compute the intensity of technology occupations in four-digit NAICS codes for 2002. The national average of technology-oriented occupations for all industries is 4.9 percent. An industry is defined high-tech if its employment share in technology-oriented occupations is at least twice this national average. As a result, Hecker defined a total of 46 four-digit NAICS codes, which were further distinguished into three levels of high-tech (Level I to III).

I only use those industry sectors that Hecker defined as so-called Level I high-tech industries. Fourteen NAICS-based sectors qualified as Level I because technology-oriented occupations accounted for "a proportion that was at least five times the average or greater and constituted 24.7 percent or more of industry employment" (Hecker, 2005, p.58). In general, these fourteen sectors group broadly into biotechnology, information technology, high-tech manufacturing, and high-tech services and R&D. We adjusted the NAICS codes to reflect the redefinition of the classification system that occurred in 2002 (see table 6.7 in appendix 6.B for a complete list of the relevant NAICS codes).

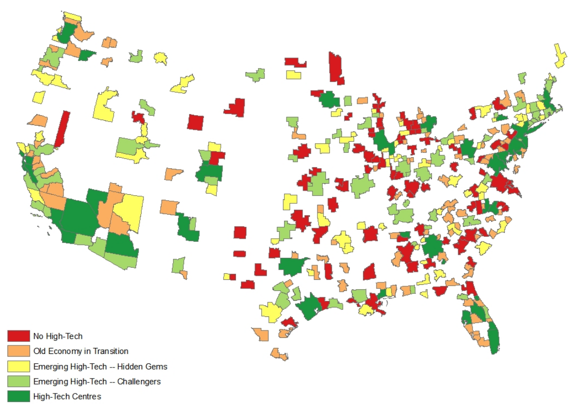
6.6.3 Towards a typology of high-tech regions

Using a similar empirical logic when determining growth-based clusters, this section outlines a methodology that identifies different types of high-technology regions. I examine the ways in which smaller metropolitan areas emerge as high-technology locations from a broad variety of data on different aspects of high-technology development, including employment, talent, entrepreneurship, and innovation. Overall, we use some twenty socio-economic variables with which I perform a principal component analysis (PCA) and a model-based cluster analysis in order to develop a typology of high-tech regions. A complete description of the different steps involved in our quantitative analysis and some specific computational results are relegated to appendix 6.A, whereas the full dataset is described in table 6.6 in appendix 6.B. From this data, it is possible to identify the following five types of regions:

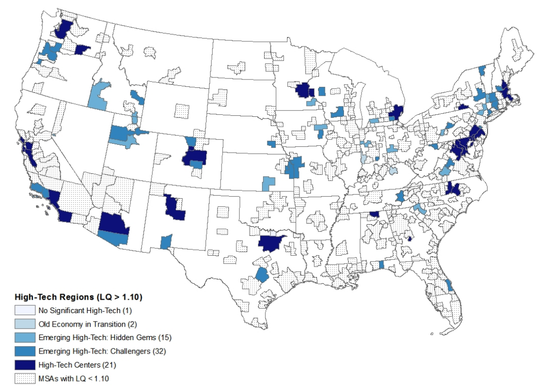
The first type of region is called *High-Technology Center* and groups regions like Silicon Valley, Boston, Washington D.C., and Seattle. These high-technology centers are highly specialized in high-technology industries and ahead of all other regions in measures of innovation, entrepreneurship, and talent. High-technology centers are hard to imitate and policymakers are better advised to learn from other locations, namely emerging high-technology regions. This preliminary analysis finds that there are two different kinds of emerging high-technology regions, which we named *High-Tech Challengers* and *Hidden Gems*. High-tech challenger regions are becoming more specialized in high-technology, they are more innovative and entrepreneurial than the average metropolitan region and they attract and retain talent. Regions that are grouped as hidden gems are unique as they show signs of improving high-tech industry activity. More than half of all metropolitan areas in the United States have, however, not successfully developed high-technology industries. In this category, 99 metropolitan areas are struggling to restructure their economies and subsequently characterised as *Old Economy Regions in Transition*. Another 85 metropolitan areas belong to a group of regions that do not have any significant high-tech industry activity and is thus labelled *Regions with No Significant High Tech*.

The main intuition behind this approach of constructing a typology of U.S. metropolitan areas is twofold. First, metropolitan areas are clustered into groups that share similar features and, second, by examining the way in which these clusters vary, we obtain additional insights into the different economic dynamics that influence each region. In other words, this method allows different metropolitan

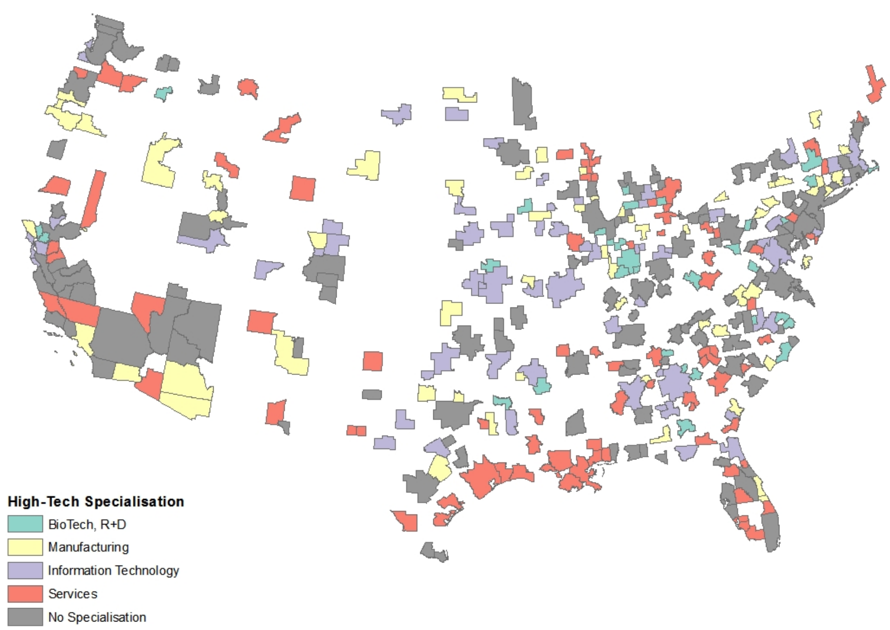
Figure 6.11: High-tech regions by type and degree of specialisation



(a) Cluster types



(b) Location quotients larger than 1.1 by type



(c) Specialisation of high-tech clusters

Notes: Specialisation is calculated on the basis of the share of high-tech employment per six-digit industry sector. If any sector's employment share exceeds 45% of total high-tech employment, the metropolitan area is classified as being specialised in this industry. Specialisation types are defined in column (4) of table 6.7. See also table 6.4 for statistics on specialisation. *Source*: Authors' calculations, County Business Patterns areas to be grouped on the basis of similar characteristics. For example, high-tech challenger regions are generally more innovative and entrepreneurial than the metropolitan average. However, they differ from high-tech centers with regard to other measures, for instance their growth dynamics. Similarly, metropolitan areas fall into the hidden gem regions category, because as a group they show improving measures of high-tech specialization and growth in industry R&D funding. Thus clustering identifies those metropolitan areas that are similar to each other as measured by the variables that reflect of innovation, talent, and entrepreneurship.

Contrary to simply ranking MSAs by a small set of variables (such as the number of high-tech jobs, location quotients, or firms), the additional information that is contained in a much broader set of variables can be utilised by reducing the dimensionality of the data via PCA. This then permits to form model-based clusters which – in contrast to other clustering or ranking techniques – are not based on arbitrary cut-off points. Since the identification of emerging high-tech regions is of greatest interest to policy makers, this analysis allows to go beyond the largest high-tech centers and examine other regions in more depth. A simple ranking along two or three single measures typically produces very unstable rankings which tend to favor the largest high-tech regions without shedding light on smaller regions with similar growth dynamics. As a result, a metropolitan region which – if only measured by its location quotient or high-tech employment – appears as “low-tech”, might nonetheless be categorized as “high-tech”. This is because all of its other characteristics are very similar to more obvious members of that cluster. The reverse is of course also true for prima facie “high-tech” candidates that end up categorized in “low-tech” clusters.

For example, although Gulfport-Biloxi, Mississippi, has a high-tech location quotient of 0.63 and high-tech employment stands at 2,732, it is grouped among the high-tech centers. In contrast, nearby Mobile, Alabama, is listed as a region with no high-tech activity. Yet, Mobile has more employees (5,709) and a slightly higher location quotient (0.74). The growth dynamics and high-tech potential in the two regions, however, differ in significant ways: While both MSAs enjoyed above average annual GDP growth of 3.9% and 5.2% from 2000 to 2005 respectively, the high-tech employment in Gulfport-Biloxi increased by almost two thirds as opposed to only one quarter in Mobile over the same time period. As a result, Gulfport-Biloxi’s high-tech specialization more than doubled compared to a modest one-sixth increase in Mobile. This deepened specialization has also materialized in higher value creation in Gulfport-Biloxi as its GDP of $97,369 per job clearly exceeds the $80,815 per job in Mobile. Furthermore, the two MSAs also diverge fundamentally in terms of their potential for innovation and entrepreneurship: R&D employment in Gulfport-Biloxi accounts for 78% of all high-tech employment, whereas a mere 2% of high-tech employment in Mobile is focused on R&D. Similarly, cumulative industry R&D funding in Gulfport-Biloxi dwarfs that of Mobile by a factor of over 20, having more than quadrupled between 2000 and 2005, when funding barely grew by half in Mobile. Lastly, Gulfport-Biloxi's self-employment is almost 40% higher than the same indicator for entrepreneurship in Mobile.

In contrast to other studies of high-technology growth which often focus only on the largest cities and regions, this chapter examines the universe of all 358 metropolitan areas in the United States (MSAs). Such a comprehensive approach allows not only to focus on the largest and most prominent high-technology regions but also on regions that may have been able to make inroads into the knowledge economy from a less advantageous position. This avoid a bias towards the largest and most successful regions.

The data covers all 358 MSAs across the contiguous United States. A focus on MSAs does not take into account that individual close-by metropolitan areas have important economic connections (through commuter ties, but also possibly through business networks such as supplier or buyer linkages). For example, the initial set of metropolitan areas shows Durham and Raleigh as separate Metropolitan Statistical Areas (MSAs). The conventional view, however, is that both are considered important parts of the Research Triangle region and we decided to combine these two regions. The top ten metropolitan areas for each cluster type are ranked in table 6.5 below.

Roughly one in ten of all MSAs (36) are identified as high-tech centers. 68 regions are classified as high-tech challengers and 71 regions are hidden gem regions. 85 regions are old economy regions in transition. And 97 regions are regions with no significant high-tech activity. Table 6.3 summarises the most important structural characteristics of these five types of clusters.[4] The map in panel 6.11(a) illustrates the spatial distribution of the different cluster types and panel 6.11(b) highlights those MSAs that have a location quotient above 1.10. Map 6.11(c) shows the industrial specialisation for each high-tech region. The relationship between sectoral specialisation, industrial concentration and differences in regional productivity are explore in more detail in section 6.6.4 below.

6.6.4 Relative productivity and the gains from specialisation

Emerging high-tech centers, in particular the smaller hidden gems, display a higher degree of high-tech specialization than other regions. Specifically, emerging high-tech centers tend to have specialized in

[4] Supplemental figures with boxplots for each group of variables across the clusters are shown in appendix 6.C.

Table 6.3: Comparison of regional characteristics

	All MSAs	High-Tech Centers	Emerging HT Challengers	Emerging HT Hidden Gems	Old Economy in Transition	No High-Tech
Industry						
Average high-tech employment (2005)	15,025	95,305	14,711	5,397	5,003	1,555
% change HT empl. (1998-2005)	40%	16%	92%	29%	40%	20%
% change HT est. (2005)	57%	46%	103%	47%	-5%	8%
Average GDP growth (%, 2000-05)	3.03%	3.73%	3.76%	3.05%	1.93%	3.23%
MSA GDP per capita ($, 2005)	$36,199	$46,005	$40,432	$34,150	$37,422	$30,147
HT location quotient (2005)	0.78	1.57	1.23	0.81	0.56	0.37
% change LQ (1998-2005)	46%	15%	82%	33%	51%	23%
Innovation						
Share of R&D empl. (2005)	12%	43%	22%	11%	4%	1%
Industry R&D funding ($mn, 2000-05)	458,467	314,052	80,798	31,871	24,850	6,896
Uni R&D expenditures ($mn, 1997-2000)	115,081	56,125	39,667	6,376	12,702	211
Patents (per 1,000 people, 1990-99)	110	476	70	89	69	53
SBIR grants (per 1,000 people, 2000-05)	6	13	8	11	1	1
Talent						
Share with bachelor degree (2005)	16.8%	20.8%	21.1%	16.6%	15.3%	13.7%
Empl. share of “creative class” (2005)	22.5%	28.5%	25.6%	22.2%	21.9%	19.0%
Research I universities (2008)	90	43	31	3	13	0
Entrepreneurship						
Avg. HT firm births (1997-2000)	29	166	25	11	14	7
% self-employed (2005)	12.5%	12.2%	13.9%	12.9%	12.8%	11.1%
Firm births (2000-05)	10,412	55,606	9,445	4,096	5,558	3,339
Firm deaths (2000-05)	9,457	49,845	8,647	3,694	5,243	3,079
VC deals (per 1,000 p., 2000-05)	4	12	5	3	2	2
Number of observations	358	36	68	71	85	97

Note: The four categories (industry, innovation, talent and entrepreneurship) mirror the logic of grouping the variables in section 6.3. The sources for this data are described in appendix 6.B.

manufacturing and service-intensive high-tech industries. Indeed, 40% of the high-tech hidden gems specialize in high-tech-manufacturing and 27% of that same group has a service-based or an R&D focus. Similarly, for the cluster of emerging high-tech challengers almost a third are specialized in manufacturing and a quarter have their strategic concentration in service-based high-tech. By contrast, almost two thirds of the high-tech centers have no specific industrial specialization. This strong specialization difference between emerging high-tech and the rest is also reflected in table 6.5 which shows that none of the top ten high-tech centers are specialized, whereas the majority of the top 10 emerging high-tech regions show a specific concentration in one of the four high-tech industry types.

Table 6.4: High-tech productivity and specialisation

	All MSAs	High-Tech Centers	Emerging HT Challengers	Emerging HT Hidden Gems	Old Economy in Transition	No High-Tech
Economic performance						
High-tech output share	11%	16%	13%	10%	9%	12%
High-tech employment share	4%	8%	6%	4%	2%	3%
Labor productivity (%)	2.9%	2.0%	2.1%	2.6%	4.9%	4.4%
MSA GDP per captia ($ nominal, 2005)	$36,199	$46,005	$40,432	$34,150	$30,147	$37,422
MSA GDP per job ($ nominal, 2005)	96,337	115,497	103,849	90,845	87,751	94,757
MSA GDP per HT job ($ nominal, 2005)	431,403	301,762	322,823	426,960	470,696	502,715
*Specialisation**						
Biotech	2%	0%	3%	7%	0%	0%
Manufacturing	54%	10%	31%	40%	50%	100%
IT	6%	10%	13%	13%	0%	0%
Services	12%	18%	25%	27%	0%	0%
No specialization	26%	62%	28%	13%	50%	0%
Number of observations	358	36	68	71	85	97

Note: * Specialisation is calculated on the basis of the share of high-tech employment per six-digit industry sector. If any sector's employment share exceeds 45% of total high-tech employment, the metropolitan area is classified as being specialised in this industry. Specialisation types are defined in column (4) of table 6.7.

While the high-tech industry has above-average (labour) productivity that is independent of its geographic location, the absolute level of productivity is highest in regions that are not considered high-tech centers, suggesting diminishing marginal productivity associated with high-tech specialization. This implies that the importance of high-technology industries to a region's productivity is the highest in regions that are not considered high-tech centers. Focusing on high-tech economic development pays off more in emerging high-tech regions, old economy regions, and regions with no significant high-tech activity than in regions that are already high-tech centers. High-tech industries' national share of output is almost three times its national share of employment and, at the national level; the GDP per high-tech job is close to 4.5 times that of an average job. This indicates that high-technology industries are critically important to the prosperity of the nation's metropolitan economies.

As table 6.4 indicates, the contribution of high-tech industries to productivity is the highest in those regions that do not have a large and diversified high-tech base. In high-tech challenger regions, the high-tech output share is 13 percent while the high-tech employment share of the overall economy is six percent. Similarly GDP per high-tech job in high-tech challenger regions is three times higher than GDP per job in the entire economy. Hidden gem regions have a high-tech output share of 10 percent and a high-tech employment share of four percent. GDP per high-tech job in these hidden gem regions is more than 4.5 times higher than GDP per job, indicating the importance of high-tech industries.

Old economy regions in transition and regions with no significant high-tech show even higher measures of high-tech productivity. In old economy regions, high-tech output share is 9 percent, significantly below the MSA average. Their high-tech employment share is two percent. GDP per high-tech job is five times higher than GDP per job. A similar pattern emerges for those regions that do not have significant high-tech industries. These findings show that the marginal contribution of an additional high-technology job is much smaller in high-tech centers than in emerging high-tech regions. One additional high-tech jobs contributes the most in regions that are restructuring or do not have much high-tech activity.

Table 6.5: Top 10 MSAs by type of high-tech region

	High-Tech Empl. (2005)	High-Tech LQ (2005)	% Change HT Empl. (1998-2005)	% Change HT Est. (1998-2005)	Specialization
High-Tech Centers					
Washington-Arlington-Alexandria, DC-VA-MD-WV	306,271	2.64	0.23	1.41	IT
Los Angeles-Long Beach-Santa Ana, CA	302,194	1.17	-0.17	0.81	Diversified
Boston-Cambridge-Quincy, MA-NH	218,392	1.96	-0.09	0.70	Diversified
San Jose-Sunnyvale-Santa Clara, CA	197,253	4.64	-0.25	0.45	Diversified
Dallas-Fort Worth-Arlington, TX	196,248	1.61	0.16	1.42	Diversified
San Francisco-Oakland-Fremont, CA	178,160	1.89	0.14	0.73	IT
Seattle-Tacoma-Bellevue, WA	162,713	2.31	0.71	0.99	Diversified
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	155,891	1.25	0.09	1.46	Diversified
Detroit-Warren-Livonia, MI	116,309	1.31	0.32	1.12	Services
San Diego-Carlsbad-San Marcos, CA	107,306	1.87	0.22	1.34	Diversified
Emerging High-Tech: Challengers					
Austin-Round Rock, TX	68,455	2.40	0.08	1.13	Diversified
Portland-Vancouver-Beaverton, OR-WA	58,646	1.35	-0.09	1.21	Manufacturing

continued on the next page

	High-Tech Empl. (2005)	LQ High-Tech (2005)	% Change HT Empl. (1998-2005)	% Change HT Est. (1998-2005)	Specialization
Kansas City, MO-KS	49,918	1.14	-0.02	1.59	IT
Hartford-West Hartford-East Hartford, CT	38,516	1.42	0.69	0.86	Manufacturing
Tucson, AZ	33,651	2.18	0.50	1.25	Manufacturing
Salt Lake City, UT	32,989	1.32	0.05	1.43	Diversified
Bridgeport-Stamford-Norwalk, CT	31,682	1.45	0.27	0.89	Manufacturing
Oxnard-Thousand Oaks-Ventura, CA	26,872	2.03	0.63	1.12	Diversified
Palm Bay-Melbourne-Titusville, FL	26,136	2.97	0.14	1.15	Manufacturing
Springfield, MA	22,161	1.77	3.33	1.12	IT
Emerging High-Tech – Hidden Gems					
Wichita, KS	37,317	2.98	4.36	2.09	Manufacturing
Dayton, OH	27,725	1.60	0.21	0.95	IT
Poughkeepsie-Newburgh-Middletown, NY	21,691	2.16	0.75	0.97	Manufacturing
Boise City-Nampa, ID	18,969	1.76	-0.15	2.19	Manufacturing
Greenville, SC	15,422	1.16	-0.14	1.46	Services
Provo-Orem, UT	12,738	1.81	0.07	1.46	IT
Cedar Rapids, IA	10,654	1.90	-0.55	1.41	Manufacturing
Kalamazoo-Portage, MI	10,327	1.67	0.49	1.57	Services
Lynchburg, VA	5,626	1.24	-0.26	1.54	Diversified
Rochester, MN	5,256	1.18	-0.51	1.26	Manufacturing
Old Economy in Transition					
Virginia Beach-Norfolk-Newport News, VA-NC	31,569	1.04	0.22	1.54	Services
Milwaukee-Waukesha-West Allis, WI	30,582	0.79	0.04	1.27	Diversified
Richmond, VA	16,565	0.66	0.03	1.46	Diversified
Buffalo-Niagara Falls, NY	16,256	0.70	0.02	1.42	Diversified
Jacksonville, FL	16,113	0.64	0.19	1.75	IT
Tulsa, OK	13,695	0.77	-0.01	1.78	Diversified
Birmingham-Hoover, AL	13,491	0.59	0.06	1.38	Diversified
New Orleans-Metairie-Kenner, LA	13,149	0.52	-0.18	1.41	Services
Omaha-Council Bluffs, NE-IA	13,076	0.69	-0.40	1.37	IT
Syracuse, NY	12,083	0.95	-0.08	1.35	Diversified
No Significant High-Tech					
Sarasota-Bradenton-Venice, FL	6,165	0.57	-0.11	1.95	Diversified
Mobile, AL	5,709	0.74	0.25	1.36	Diversified
Spokane, WA	5,427	0.63	-0.40	1.44	Diversified
Bakersfield, CA	5,393	0.64	0.21	2.30	Services
El Paso, TX	4,731	0.49	0.13	2.04	Diversified
Cape Coral-Fort Myers, FL	4,609	0.52	0.43	1.81	Diversified
York-Hanover, PA	3,856	0.49	-0.11	1.63	Services
Scranton–Wilkes-Barre, PA	3,687	0.33	-0.54	1.80	Diversified
Reading, PA	3,537	0.49	-0.41	1.48	Services
Vallejo-Fairfield, CA	3,438	0.66	0.45	1.81	Biotech

6.7 Summary and future work

In this short paper I am proposing an alternative classification of MSAs on the basis of clusters that were formed by principal component analysis from economic variables that are relevant for regional growth.

These variables include labour productivity growth, measures of local industry mix, human capital, entrepreneurship and innovation. The usefulness of the classification of MSAs in to these new regional groups was then tested by testing whether the presence of localised spillovers from university R&D and "creative people" to new high-tech firm formation varies regionally. For this purpose, I used the growth-based regional clusters to control for the presence of cluster-specific fixed-effects in the context of firm's location choices.

The second part of this chapter contains the first elements for the analysis of the main drivers of regional specialisation and industrial concentration of urban economies. Future work would relate these structural differences of metropolitan areas to specific divergences in the economic performance of cities with the aim to enrich recent work on the urban business cycle (e.g. Owyang, Piger, Wall, and Wheeler, 2008).

6.A Model-based clustering

To examine the ways in which metropolitan areas differ in their high-tech development, we employ a principal component and model-based cluster analysis to create typologies of high-tech regions. The main idea behind a regional typology of US metropolitan areas is to cluster observations into groups that share similar features and then investigate the way in which the groups differ. Clustering methods are among the most widely used techniques to partition such data into meaningful sub-groups. However, most clustering approaches are largely heuristic and lack the rigor usually associated with structural models. In order to address these issues, we use a procedure that adopts a statistical model to group the data. Specifically, we follow Fraley and Raftery (2002, 2006) who implement a cluster analysis based on parameterized Gaussian mixture models.

A distinctive advantage of a model-based clustering approach is that it permits the use of more objective model selection techniques (such as the Bayesian Information Criterion [BIC]) to compare outcomes rather than the arbitrary choices of more popular clustering approaches such as k-means or hierarchical clustering. Model-based clustering thus provides us with a systematic means of selecting both the parameterization of the model and also the number of clusters.

In developing the typology, we use a wide range of data that reflect the MSAs economic performance and their associated degrees of industrial specialization as well as input-based characteristics, such as quality of the labor force, R&D funding and employment concentration. The 20 individual variables used for the cluster analysis are listed in table 6.6 in appendix 6.B. The data can broadly be grouped into four thematic categories, namely, economic performance, talent, innovation, and entrepreneurship.

Our empirical strategy can thus roughly be described in three steps: First, we perform an exploratory investigation of spatial properties of the data. Second, since cluster analysis data usually consists of independent multivariate observations, the dimensions of the data set are reduced by means of a principal component analysis (PCA) in order to derive independent key drivers. Third, the principal components form the basis for the model-based cluster analysis with the aim to create different groups of MSAs.

From our total of twenty socio-economic variables, the principal component analysis identifies six factors with eigenvalues larger than unity with explain just over 75 percent of the total variation in the data. We include a measure of natural amenities in order to control for the regional variation in the quality of life among MSAs. The model-based clustering procedure and the BIC outcome lead to the selection of five clusters of MSAs (BIC = -8,325.8) using the VVI parameterization (diagonal, varying volume, varying shape) for the component covariance matrix.

Once we determined the five clusters, we illustrate dynamics of growth for each type (i.e. industry measures such as high-tech employment, number of establishments, gross domestic product, as well as measures of innovation, talent, and entrepreneurship). These measures describe the five different regional types and were ultimately instrumental in providing a qualitative characterization of each cluster; “High-tech Centers”, “Emerging High-tech: Challengers”, “Emerging High-tech: Hidden Gems”, “Old Economy in Transition” and “No High-tech”.

6.B Data sources and definitions

Table 6.6: Variables used for model-based cluster analysis

Variable	*Description*	*Source*
GDP per capita (2005)	MSA per capita GDP	BEA-REA
Average GDP growth (2000-2005)	Annualized compound average GDP growth	BEA-REA
Share of GDP from High-Tech (2005)	Percentage of metropolitan GDP due to high-tech industries using 3-digit NAICS codes	BEA-REA
Location Quotient for High Tech	Ratio of share of MSA High-Tech employment to share of national high-tech employment. The definition of high-tech employment contains 14 industries at the NAICS 4-digit level.	Hecker (2005), BLS
Change in Specialisation	% Change LQ High-Tech from 1998 to 2005	
Firm Births/Deaths (1998-2004)	Cumulative number of firm births to firm deaths from 1998 to 2004	SBA-OA

continued on the next page

Variable	*Description*	*Source*
High-Tech Firm Births (1997-2000)	Number of new high-tech firm formations	Data set obtained from Woodward, Figueiredo, and Guimarães (2006, WFG hereafter)
University R&D Expenditures (1991-1997)	Yearly total university R&D expenditures in science and engineering for Doctoral Universities (I and II), Schools of Engineering, and Technology and Research Universities (I and II), according to the Carnegie classification	Data set obtained from WFG
Research Universities	Number of Research Universities per MSA	CCIHE 2000
Share of R&D Employment (2005)	MSA-share of employment in R&D-intensive industries (NAICS code 5471, also referred to as Scientific R&D Services).	CBP (1998-2005)
Cumulative R&D Funding (1998-2005)	Cumulative amount of non-university NSF funding in nominal dollars from 1998 to 2005. State level data attributed to MSA-level using the MSA-share of employment in R&D-intensive industries (NAICS code 5471, also referred to as Scientific R&D Services) as attribution key for state-level data.	NSF-SRS (1998-2005)
Number of Patents (1990-1999)	Cumulative number of MSA-level patent registrations from 1990 to 1999	USPTO
SBIR Grants (1998-2004)	Total number of state level data attributed to MSA-level using the MSA-share of employment in R&D-intensive industries (NAICS code 5471, also referred to as Scientific R&D Services) as attribution key for state-level data.	SBA-OT
VC Deals per Capita (2000-2005)	Number of venture capital deals per capita where state level data is attributed to the MSA-level by using population weights.	SBA-OA
% Creative Employment (2000)	Creative class measure with percent employed in creative class occupations and a metro/nonmetro indicator for all counties, 1990 and 2000.	USDA-ERS
% with Bachelors Degree (2000)	Population 25 years and over with Bachelor's degree as percentage of total population segment. (Series P037001, P037015, P037032).	CPS
% Self-Employed (2000)	Percentage of households with self-employment income (Series P052001 and P060002).	CPS
Employment Concentration HHI	Herfindahl-Hirschman Index of employment concentration by large firms (> 1000 employees).	CBP (1998-2005)
Natural Amenities Scale	Measure of the physical characteristics of a county area that enhance the location as a place to live combining six measures of climate, typography, and water area that reflect environmental qualities most people prefer.	USDA-ERS
Employment in High-Tech (2005)	Number of employees in high-tech industries as defined in table 6.7.	CBP (1998-2005)
Employment: % Change (1998-2005)	Absolute percentage change of number of employees from 1998 to 2005.	CBP (1998-2005)
High-Tech Establishments (2005)	Number of establishments in high-tech industries	CBP (1998-2005)

Notes: BEA-REA: Regional Economic Accounts from the Bureau of Economic Analysis; CBP: Census Bureau County Business Patterns; CCIHE: Carnegie Classification of Institutions of Higher Education; CPS: 2000 Decennial Current Population Survey, U.S. Census Bureau; NSF-SRS: U.S. National Science Foundation, SRS Publications and Data, Federal Funds for R&D Series; SBA-OA: U.S. Small Business Administration, Office of Advocacy; SBA-OT: U.S. Small Business Administration, Office of Technology (Small Business Innovation and Research/STTR); USDA-ERS: U.S. Department of Agriculture Economic Research Service; USPTO: U.S. Patent and Trademark Office.

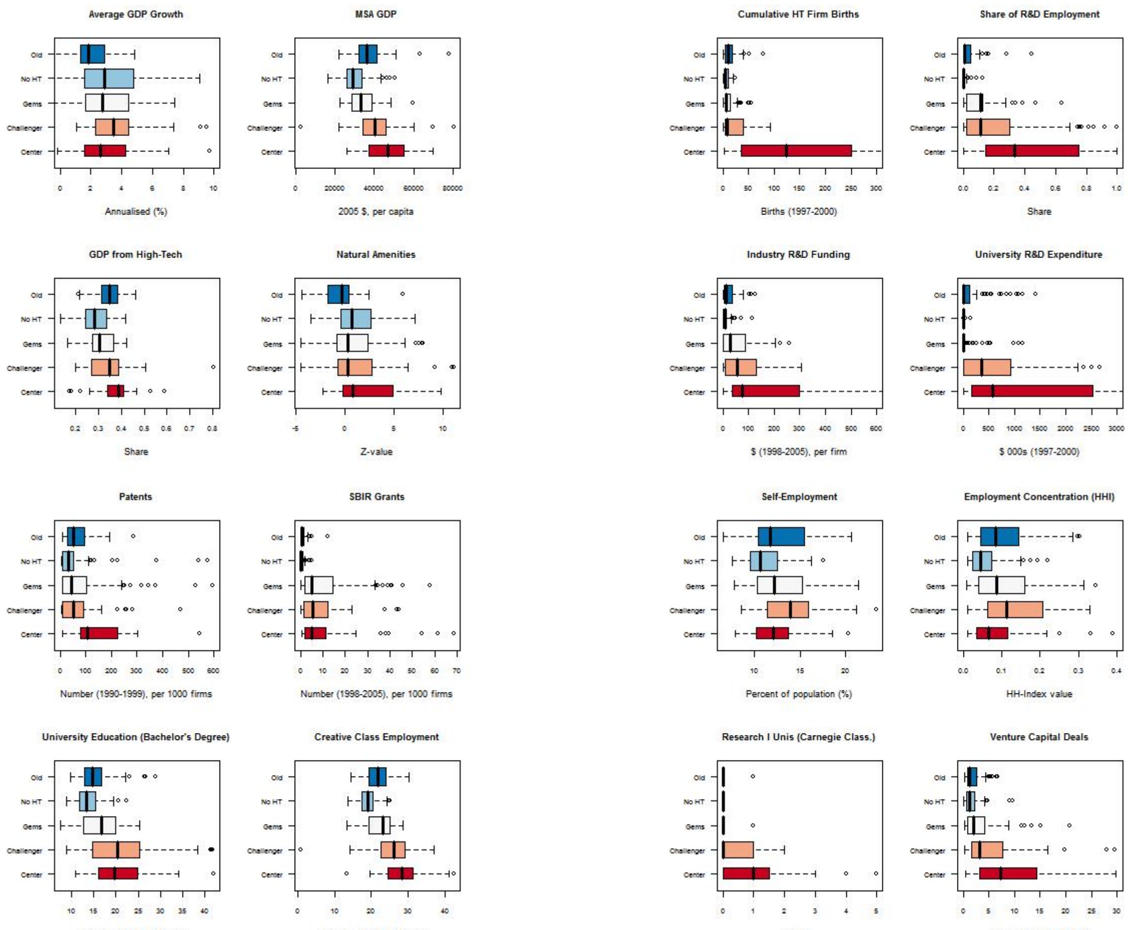
Table 6.7: High-Tech NAICS classification (level I industries)

1997 NAICS (1)	*2002 NAICS* (2)	*Industry Description* (3)	*Type* (4)
3254		Pharmaceutical and Medicine Manufacturing	Biotech
3341		Computer and Peripheral Equipment Manufacturing	
3342		Communications Equipment Manufacturing	Manufacturing
3344		Semiconductor and Other Electronic Component Manufacturing	
3345		Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	
3364		Aerospace Product and Parts Manufacturing	
5112		Software Publishers	
5110 (part)	5161	Internet Publishing and Broadcasting	
513390	5179	Other Telecommunications	Information Technology
514191, 514199	5181	Internet Service Providers and Web Search Portals	
514210	5182	Data Processing, Hosting, and Related Services	
5413		Architectural, Engineering, and Related Services	
5415		Computer Systems Design and Related Services	Services and R&D
5417		Scientific Research and Development Services	

Source: Authors' classifications of types, Hecker (2005); US Census Bureau for the correspondence between 1997 NAICS and 2002 NAICS.

6.C Supplemental figures

Figure 6.12: High-tech cluster characteristics



The boxplots show the distribution of key variables across the five types of US metropolitan high-tech industry clusters. The clusters were identified using a model-based cluster analysis. *Source*: County Business Patterns, author's calculations.

6.D References

AEA (2003): *Defining the High-Tech Industry: AeA's New NAICS-based Industry Classification*. American Electronics Association, Washington, D.C.

AIGINGER, K., AND S. W. DAVIES (2002): "Industrial Specialisation and Geographic Concentration: Two Sides of the Same Coin? Not for the European Union," *Journal of Applied Economics*, 7(2), 231–248.

AIGINGER, K., AND E. ROSSI-HANSBERG (2006): "Specialization and Concentration: A Note on Theory and Evidence," *Empirica*, 33(3), 255–266.

CHINITZ, B. (1961): "Contrasts in Agglomeration: New York and Pittsburgh," *American Economic Review*, 51(2), 279–289.

CRONE, T. M. (2005): "An Alternative Definition of Economic Regions in the United States Based on Similarities in the State Business Cycle," *Review of Economics and Statistics*, 87(4), 617–626.

CRONE, T. M., AND A. CLAYTON-MATTHEWS (2005): "Consistent Indices for the 50 States," *Review of Economics and Statistics*, 87(4), 593–603.

CUTRINI, E. (2009): "Using Entropy Measures to Disentangle Regional from National Localization Patterns," *Regional Science and Urban Economics*, 39(2), 243–250.

DURANTON, G. (2007): "Urban Evolutions: The Fast, the Slow, and the Still," *American Economic Review*, 97(1), 197–221.

DURANTON, G., AND D. PUGA (2000): "Diversity and Specialisation in Cities: Why, Where and When Does it Matter?," *Urban Studies*, 33(3), 533–555.

——— (2001): "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," *American Economic Review*, 91(5), 1453–1477.

FRALEY, C., AND A. E. RAFTERY (2002): "Model-Based Clustering, Discriminant Analysis and Density Estimation," *Journal of the American Statistical Association*, 97(458), 611–631.

——— (2006): "MCLUST Version 3 for R: Normal Mixture Modeling and Model-Based Clustering," Technical Report 504, University of Washington, Department of Statistics.

GLAESER, E. L. (1999): "Learning in Cities," *Journal of Urban Economics*, 46(2), 254–277.

HECKER, D. E. (2005): "High-technology employment: a NAICS-based update," *Monthly Labor Review*, 128(7), 57–70.

HENDERSON, J. V. (1974): "The Sizes and Types of Cities," *American Economic Review*, 64(4), 640–656.

LEHTO, E. (2007): "Regional Impact of Research and Development on Productivity," *Regional Studies*, 41(5), 623?–638.

MAYER, H. (2009): "Bootstrapping High-Tech: Evidence from Three Emerging High Technology Metropolitan Areas," Metropolitan Economy Initiative No. 10, Brookings Institution, Washington, DC.

O'HUALLACHÁIN, B. (2008): "Regional Growth Transition Clubs in the United States," *Papers in Regional Science*, 87(1), 33–54.

O'SULLIVAN, D., AND D. UNWIN (2003): *Geographic Information Analysis*. John Wiley and Sons, Inc., Hoboken, NJ.

OWYANG, M. T., J. M. PIGER, H. J. WALL, AND C. H. WHEELER (2008): “The Economic Performance of Cities: A Markov-switching Approach,” *Journal of Urban Economics*, 64(4), 538–550.

OWYANG, M. T., D. E. RAPACH, AND H. J. WALL (2009): “States and the Business Cycle,” *Journal of Urban Economics*, 65(1), 181–194.

PANEK, S. D., F. T. BAUMGARDNER, AND M. J. MCCORMICK (2007): “Introducing New Measures of the Metropolitan Economy: Prototype GDP-by-Metropolitan-Area Estimates for 2001?-2005,” *BEA Survey of Current Business*, 87(11), 79–113.

PORTER, M. E. (2000): “Location, Competition, and Economic Development: Local Clusters in a Global Economy,” *Economic Development Quarterly*, 14(1), 15–34.

ROSSI-HANSBERG, E. (2005): “A Spatial Theory of Trade,” *American Economic Review*, 95(5), 1464–1491.

SAXENIAN, A. (1994): *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press, Cambridge, MA.

SCHERER, F. M. (2000): “Professor Sutton’s ‘Technology and Market Structure’,” *Journal of Industrial Economics*, 48(2), 215–223.

SUTTON, J. (2001): *Technology and Market Structure: Theory and History*. MIT Press, Cambridge, MA, 2nd edn.

WOODWARD, D., O. FIGUEIREDO, AND P. GUIMARÃES (2006): “Beyond the Silicon Valley: University R&D and High-Technology Location,” *Journal of Urban Economics*, 60(2), 15–32.

Chapter 7

Towards a generalisation of the representative agent model for quality of life

7.1 Abstract

Given its relatively restrictive set of assumptions, this chapter proposes the generalisation of two important aspects of the standard Rosen-Roback model with regard to quality of life. First, I examine the impact of household heterogeneity on hedonic measures of quality of life. In particular, I outline how income heterogeneity could be accommodated by adjusting the traditional model. I use a quantile framework to identify evidence of income-related variation in the willingness to pay for amenities. Second, I discuss the implication of relaxing the fixed labour supply assumption on quality of life. In standard economic theory, labour supply decisions depend on a complete set of prices, not just wages. Local variation in non-wage prices can thus have an impact on households' labour supply decisions. Endogenous labour supply decisions introduce an important link between local amenities and differences in local labour markets. This chapter also investigates to what extent evidence of local differences in the average weekly hours worked is inconsistent with the assumption of fixing labour supply as a constant in the standard urban location model by Roback (1980, 1982). I examine empirical implications for hedonic estimates of quality of life, focusing on previous work that uses microdata while maintaining the fixed labour supply assumption.

Keywords: Rosen-Roback model, quality of life, generalised representative agent model, heterogeneity, labour supply
JEL Codes: E01, H41, Q51, R23

7.2 Introduction

The restrictive assumption of a homogeneous labour force with workers that display homothetic preferences is one of the most obvious draw-backs of the canonical framework of compensating differentials due to Rosen (1979) and Roback (1980, 1982). In a later paper, Roback (1988) relaxes the homogeneity assumption by introducing two types of workers (high skill and low skill). Similar work by Beeson (1991) examines the relative importance of demand- versus supply-driven explanations of local differences in the returns to worker characteristics, highlighting that equilibrium differences in these returns can exist if the characteristics affect the demand for housing relative to amenities. In the standard setting of the canonical Rosen-Roback model, however, all households face the same set of implicit amenity prices since the willingness to pay for amenities cannot vary across households. Yet theory suggests that differences in household characteristics influence the marginal willingness to pay for most goods. It is of particular relevance for public policy to identify the sources of this variation vis-à-vis the most common dimensions of household heterogeneity, namely income and age (e.g. Evans and Schaur, 2010).

The remainder of this chpater proceeds as follows. Section 7.3 outlines the logic of introducing heterogeneity to the Rosen-Roback models and section 7.4 provides a first set of empirical evidence that is consistent with income and age heterogeneity among households. Section 7.5 turns to the second generalisation of the model by endogenising labour supply and 7.6 proposes an alternative measure for implicit amenity prices in terms of observable quantities that can be estimated via hedonic methods. Section 7.7 discusses the consequences of elastic labour supply for hedonic quality-of-life estimates, particularly focusing on potential inconsistencies in previous work that uses microdata while maintaining the fixed labour supply assumption. Section 7.8 concludes this chapter with a short summary of future research on this topic.

7.3 Quality of life and household heterogeneity

In this section, I trace the first steps to show that income heterogeneity can be accommodated within the traditional Rosen-Roback model. This motivates the use of a quantile regression approach to empirically account for heterogeneity in the context of a first set of hedonic quality-of-life estimates in section 7.4. I begin by assuming that utility is additively separable from leisure such that the utility function corresponding to equation (A-1) from chapter 3 takes the following form:

$$U(x,h;q) = \left[\theta q^{\rho} + (1-\theta) y^{\rho}\right]^{1/\rho}, \tag{7.3.1}$$

where the subutility function for the composite commodity that includes the traded good and housing services, $y = \psi(x,h)$, is homothetic and the elasticity of substitution between the quality of life q and y is given by $\sigma = \frac{1}{1-\rho}$. The inverse demand function for the quality of life, $p^q(q)$, can be derived directly from equation (7.3.1) by invoking Wold's theorem and using the Euler aggregation which yields:[1]

$$p^q = \theta m q^{\rho-1}/U^{\rho}. \tag{7.3.2}$$

In the extreme case where quality of life and income are perfect substitutes (i.e. $\rho = 1$ and $\sigma \to \infty$), the implicit price for the quality-of-life is independent of income as equation (7.3.2) reduces to $p^q = \theta$. However, with unitary elasticity of substitution (i.e. $\rho = 0$ and $\sigma = 1$), household utility functions are Cobb-Douglas which means that $p^q = \theta m/q$. The willingness to pay for quality of life is now positively related to the level of income. Because p^q is a linear combination of the individual implicit amenity prices p^{a_k}, Cobb-Douglas utility also implies that at least some of the equilibrium implicit amenity prices depend on the level of income. Thus, unless quality of life and income have a one-for-one trade-off, estimating implicit prices with a standard conditional mean model might not be fully consistent with the underlying utility maximisation set-up of the Roback-HBB model. In view of this, I explore potential variation in the willingness to pay from income (i.e. wage and housing expenditure) heterogeneity by using a quantile regression approach in section 7.4.

7.4 Results from quantile regressions

The true complexity of local housing and labour markets is likely to produce attribute prices that vary across the population which violates the common assumption of homogenous households. This type of misspecification can lead to estimated average effects that might vary widely across different specifica-

[1] In general, Wold's theorem states that $\boldsymbol{p}/m = \nabla U/\nabla U\boldsymbol{X}$ where $\boldsymbol{X}$ and $\boldsymbol{p}$ are vectors of goods and prices and m is income. For homogeneous utility functions, the Euler theorem states that $\nabla U\boldsymbol{X} = U$ which implies that an individual price is simply $p_i = m\frac{\partial U}{\partial x_i}/U$.

tions. In the context of explaining variations in the urban wage premium, an emerging strand of literature has begun to relax the strong assumption of homotheticity of households' preferences. In these models, heterogeneous workers with non-homothetic preferences give rise to local variations in the returns to education that may vary with location-specific amenities. Indeed, with non-homothetic preferences, location-specific variation in productivity and amenities on the returns to schooling (Black, Kolesnikova, and Taylor, 2009a; Lee, 2010).

In section 7.3, I argued that income-related household heterogeneity can be accommodated within the Roback-HBB framework if household utility is Cobb-Douglas. When I derive a set of implicit amenity prices by estimating the quality-of-life model using a quantile hedonic price function, I am able to uncover substantial variation in the hedonic prices for a range of amenities.[2] This is consistent with variation in the willingness to pay for certain amenities across the distribution of rents and wages. This variation is best made visible by plotting estimates of the hedonic gradients, dr^j/da_k^j and dw^j/da_k^j, across the quantiles of the housing expenditure and wage distributions.

7.4.1 Income heterogeneity

Figures 7.1(a) and (b) show such disparities in the willingness to pay estimates for most of the sixteen BBH amenities across the sample of metropolitan counties. While the parameter estimates from the housing regression generally show higher variation across quantiles than those from the wage regression, the parameters tends to be either increasing or decreasing in direction of the quantiles. In some instances, this also means that the sign of the hedonic gradient changes across quantiles. In combination, these patterns not only imply that the willingness to pay (and hence the implicit price) can increase or decreases with (wage) income and housing expenditure, but also that – depending on the amenity – that the price might be positive or negative across different quantiles. For example, figure 7.1(a) reveals different signs between lower and higher quantiles on the parameters for the number of heating degree days, the teacher-pupil ratio and TSD facilities. Estimating average effects might thus paint only an incomplete picture of the true nature of quality-of-life difference across locations as the hedonic estimates are influenced by the skewed nature of the wage and housing expenditure distribution.

Figure 7.1(c) shows the distributional disparities of the quality of life indices for the 10% and for the

[2] See e.g. McMillen (2008) and Redfearn (2009) for a more detailed discussion of hedonic quantile regressions and applications that deal with hedonic attributes that vary spatially and temporally.

90%. As a direct consequence, of the variation in amenity valuation across incomes, our quantile estimations also reveal (not shown here), that the means that, among otherwise identical households, wealthier households (90%-ile housing expenditure) prefer larger cities compared to poorer households (10%-ile, not shown here). This evidence is corroborated further, by examining the differences in quantile-based hedonic valuations of specific consumption amenities that tend to favour larger cities. Table 7.1 reports the conditional mean and quantile estimates of full implicit prices for a number of popular consumption and cultural amenities, namely restaurants, bars, venues with live performances, such as theatres, cinemas, bowling alleys, museums, zoos and universities. With the notable exception of the implicit price for museums, all price estimates seem reasonable and consistent with intuition, even if comparing some of the levels might warrant further investigation. Judging by the estimate for the average household's willingness reported in columns (1) and (2) alone, museums appear to make a negative contribution to the quality of life. However, the quantile-based estimates do not only uncover difference in the willingness to pay for museums, but they also show that the lower quantiles view museums as a disamenity while being highly valued by the highest income quantile. Similar variations in income-dependent valuations also obtain for live performance venues and – albeit with a different sign – for bowling alleys. These results are consistent with the evidence on "consumer cities" presented in Glaeser, Kolko, and Saiz (2001).[3]

Table 7.1: Implicit prices for selected consumption amenities

	per 1,000 people		per unit of amenity				
	OLS		OLS		Quantiles		
	Val.	(s.e.)	Val.	(s.e.)	10%	90%	IQR
	(1)		(2)		(3)	(4)	(5)
Restaurant, bar	$244.7	($145.2)	$104.1	($61.8)	$55.4	$164.7	$45.5
Live performance venue	15,963.0	(2,935.7)	214.9	(39.5)	-10.1	532.0	189.0
Bowling alley	-11,053.4	(3,621.1)	-212.5	(69.6)	45.0	-497.3	-141.0
Movie theatre	12,622.0	(4,240.7)	200.1	(67.2)	106.5	316.6	87.5
Museum	-658.6	(338.9)	-14.1	(7.3)	-73.1	63.6	50.8
Zoo	13,077.2	(5,228.7)	51.5	(20.6)	27.1	80.7	22.3
University	8,667.4	(5,883.2)	110.9	(75.3)	59.0	175.5	48.5

Notes: * Based on the improved specifications "$QOLI_3$" in column (5) of in table 3.8. Estimates in columns (1) and (2) are two-stage FGLS estimates and one-stage quantile estimates for the all amenities.

[3] I also find similar effects for some of the urban public transport amenities and and a few of the climate amenities (cf. Brueckner, Thisse, and Zenou, 1998; Glaeser, Kahn, and Rappaport, 2008)

The spatial pattern of quality-of-life preferences across age and housing consumption are illustrated in figure 7.2 below. There are a number of patterns that emerge. Above all, there the quality of life rankings are appear very robust when it comes to the most desirable places and least desirable places to life. Independent of age and the level of housing expenditure, the metropolitan north-east and coastal California are some of the most desirable places to live in the nation, whereas parts of the rural Mid-West, the deep South and rural Appalachia are the least desirable. However, significant valuation differences emerge when looking across the age distribution and across the housing expenditure distribution.

7.4.2 Amenity consumption over the life-cycle

First, there are significant valuation differences in the quality-of-life across the housing expenditure distribution. Figure 7.2(a) illustrates which counties have the highest differences in ranking when comparing the lowest quantile (10%ile) with the highest quantile (90%ile). Overall, counties in the northeastern US, mid-Atlantic states, and portions of the upper Mid-West (the Rust Belt) with a relatively high degree of urbanisation are on balance much higher ranked by households with low housing expenditure. By contrast, counties across much of the South and South-West (the Sun Belt) are ranked higher by high expenditure households. Location-specific amenities are by their very nature characterised by an indivisibility which implies that households can only change the consumption of amenities by changing their location. In the context of the standard Rosen-Roback framework, wages and amenities are substitutes which implies that households can smooth their amenity consumption over their life times only by changing location. Determining the net effects of whether people move from low-amenity locations to high-amenity locations over the life cycle or vice-versa would require developing a formal model. Housing-rich but income-poor elderly homeowners might be particularly sensitive to rising tax burdens, and anecdotal evidence suggests that some move to reduce their tax burden. Indeed, Shan (2010) finds strong evidence for the link between increases in property tax levels and the mobility rate of elderly homeowners.[4]

Second, there are also significant quality-of-life ranking differences across the age distribution. In particular, we are looking at two age groups, those at the beginning of the professional life (25 to 34 years of age) and people close to retirement (ages 55 and above). A casual inspection of figure 7.2(b) reveals

[4] The life-cycle amenity consumption aspect has – to the best of our knowledge – not yet been introduced into the Rosen-Roback framework. There are obvious parallels to the treatment of job amenities and wage differentials in search models in labour economics. See, for instance, Nosal and Rupert (2007) for a model where (job) amenities are explicitly included in workers' job and wage choices over the life cycle.

a pattern which – at least in part – seems very intuitive: Young households give urban counties in the south a higher ranking and traditional retirement destinations such as Florida, Idaho or Montana tend to be preferred by households closer to retirement age. An additional, more subtle pattern may be decerned when combining the different preferences across the housing expenditure distribution with those across the age distribution as is shown in figures 7.2(c) and 7.2(d).

In contrast to the life-cycle consumption profile of non-housing goods which is hump-shaped, the consumption profile for housing has been shown to first increase monotonically and then to flatten out across all levels of housing consumption (Yang, 2009). Furthermore residential mobility tends to decrease with age, but increases with income (Bartik, Butler, and Liu, 1992). In view of these two stylised facts, it is informative to consider the ranking preferences of different housing expenditure quantiles across the two age groups. Specifically, figure 7.2(c) illustrates the ranking dissimilarity between high housing consumption retirees (rich, mobile) and low housing consumption young (poor, but mobile) on the one hand, and figure 7.2(d) shows the valuation differences for low housing consumption retirees (poor, immobile) and the highest housing consumption young (rich, mobile).

The next section now turns to the second aspect of generalisation in the standard Rosen-Roback model, discussing the implication of relaxing the fixed labour supply assumption on quality of life estimates.

7.5 Quality of life and labour supply

Standard theories in labour economics assume that households' labour supply decisions depend on a complete set of prices. Labour supply decisions are thus not affected by wages alone, but also by other prices such as housing or the prices of non-housing consumption goods. Since these non-wage prices too are subject considerable local variation, it is possible that the labour supply in two locations with identical nominal wages differs due to local price variation alone. For example, assuming that workers are equally productive in location i and location j, sufficiently large differences in the quality of life due to variation in local amenities ($q_i > q_j$) might drive up equilibrium housing prices in location i, thereby inducing workers to work more in location i. Recently, Black, Kolesnikova, and Taylor (2009b) address this issue in a simple equilibrium setting, demonstrating that the effects of wage and non-labour income on labour supply typically differ by location.

A cursory glance at the PUMS data from the 2000 Census confirms that the average hours worked vary greatly across locations even among individuals considered full-time workers.[5] Figure 7.3(a) illustrates the distribution of the average hours worked for full-time workers across the public-use microdata areas (PUMAs) within the contiguous US. The average amount of weekly hours worked varies from 41 hours to just over 50 hours across PUMAs, a non-trivial difference in the average length of the work week of more than one working day. Furthermore, this variation in the average labour supply of full-time workers is not evenly spread across space as the average hours worked per week seem to be higher in more densely populated urban areas.

Figure 7.3(b) highlights this pattern for metropolitan areas on the East Coast. The distinct peaks in weekly average working hours (dark red) for Manhattan and Stamford within the New York City MSA and for central Boston and the Cambridge area in the Boston MSA are clearly visible. This is fully consistent with the recent evidence by Rosenthal and Strange (2008) who document a positively-increasing relationship between agglomeration and hours worked. Table 7.2 illustrates that this effect is particularly pronounced among professional young male worker in metropolitan areas who work on average in excess of three hours per week more than their non-professional, non-urban counterparts.[6]

Table 7.2: Average hours worked among male full-time workers

Occupational category	*Area*	*Young Males*	*Middle-aged Males*
Non-professional workers	Non-metro	45.76	45.71
	Metro	45.29	45.29
Professional workers	Non-metro	47.86	48.16
	Metro	49.02	48.57

Notes: Full-time workers are defined using the standard BLS definition as persons who work 35 hours or more per week. Hours worked are based on the "usual hours worked per week" in the 2000 Census. Young males are individuals aged between 30 and 39 and middle-aged males are between 40 and 49 years of age. Professional workers are individuals in occupations categorised as professional in the OCCSOC codes for the 2000 Census and who have at least a bachelor's degree and more. Metro and non-metro areas are defined using the USDA ERS rural-urban continuum.

This chapter explores to what extent evidence of local differences in full-time workers' average hours worked is inconsistent with the assumption of fixing labour supply as a constant in the standard urban location model by Roback (1980, 1982). In particular, I examine the empirical implications for hedonic

[5] Full-time workers are defined using the standard BLS definition as persons who work at least 35 hours or more per week. Using a more restrictive definition of workers who work at least 35 hours a week during at least 26 weeks of the year yields virtually identical results.

[6] In addition to self-selection of hard workers into large cities and productivity-driven agglomeration externalities, Rosenthal and Strange (2008) emphasise that this behaviour is also consistent with the presence of "a kind of urban rat race".

estimates of quality of life and show that if local variations in amenities indeed induce sufficiently large differences in local labour supply responses, previous hedonic quality of life estimates may be inaccurate. Beyond improving estimates, I relax the assumption of labour supply decisions that ignore leisure which introduces an important link between local amenities and differences in local labour market behaviour. While the existence of regional and urban labour markets is well-established, explaining local differences in labour supply behaviour are usually beyond the reach of the standard paradigm in urban economics (Zenou, 2009). Interregional differences in hours of work might help to determine whether leisure is complementary with quality of life. This link between leisure and quality of life – via their mutual relation to population density – is also key to a more detailed understanding of the channels through which agglomeration economies operate.

7.6 Local variations in labour supply behaviour

In the canonical Rosen-Roback framework, households are assumed to supply one unit of labour such that labour supply is fixed across locations. While Roback (1980) relaxes this assumption with "most of the qualitative results and insights unchanged", she also acknowledges that "it may be instructive to look for interregional differences in hours of work" (Roback, 1980, p.25) as non-trivial interaction effects between local amenities and labour supply are likely.[7] Yet previous empirical work has not examined this issue in more detail. In view of the large local variations in labour supply behaviour discussed above, an extended version of the basic model with endogenous labour supply is of particular relevance for empirical applications, such as hedonic estimates of quality of life. Furthermore, relaxing the assumption of spatially inelastic labour supply highlights likely pitfalls of previous empirical estimates of households' willingness to pay for quality of life. Modelling labour supply as a choice variable, the next section extends the standard Rosen-Roback setting such that the effects of wage and non-labour income on labour supply can differ by location. I then introduce a new representation of locational equilibrium that leads to a modified version of the familiar expression of implicit amenity prices for empirical purposes.

[7] Roback's (1980) extension to accommodate flexible labour supply does not incorporate non-wage income in contrast to standard models in the labour supply literature (e.g. Burtless and Hausman, 1978; Heckman and MaCurdy, 1980). Furthermore, her modified specification remains somewhat ambiguous with regard to the precise impact of hours worked on implicit amenity prices.

7.6.1 Endogenous labour supply in the Rosen-Roback framework

In the basic Rosen-Roback model implicit amenity prices follow from wage and rent (housing expenditure) differentials in a dual-market sorting equilibrium. The national (closed) economy is characterised by a finite number of spatially-bounded localities (e.g. counties) that differ with regard to their specific combination of wages w and the cost of housing or rents r. Places can also vary in terms of the prevailing level of quality of life q, where q^j is some index of the area j-specific bundle of K amenities, $a^j_{k \in K}$, such that $q^j = \phi(\boldsymbol{A}^j)$ and $\boldsymbol{A}^j = \{a^j_1, \ldots, a^j_k\}$ for $j = 1, 2, \ldots, J$. The national economy is divided into rural and urban areas where an arbitrary metropolitan region is composed of one or several localities. Across these sites, fully mobile households maximise their well-being subject to a budget constraint and footloose firms minimise their cost by making their respective location choices.

In site j, homogenous households enjoy the local quality of life q^j and consume a composite commodity that consists of a nationally-traded good x with price p_x which will be taken as numéraire and local, non-tradable housing services h^j which cost r^j. Households are assumed to live and work in the same location where they receive income from supplying labour at the local wage rate w^j. In addition to wage income, total household income $m^j = w^j + N$ also consists of a locally invariant component of non-wage income N. Instead of assuming that each worker supplies a fixed amount of labour, workers are now assumed to allocate their total amount of time between t hours of market work and l hours of non-productive leisure, where the total amount of time per period is normalised to 1 such that $l = (1 - t)$. Temporarily dropping superscripts for convenience, the standard utility maximisation problem can then be expressed as:

$$\max \; U(x, h, l; q) \qquad \text{s.t.} \;\; w(1 - l) + N = x + rh \tag{7.6.3}$$

The demand for leisure associated with this maximisation problem is $l = l(w, r, N; q)$ and the corresponding indirect utility function is:

$$V(w, r_h, N; q) = U[x(w, r, N; q), h(w, r, N; q), l(w, r, N; q); q] \tag{7.6.4}$$

Locational equilibrium implies identical levels of well-being across locations as households cannot improve their utility by relocating. Taking the total differential of the representative household's indirect utility function V in equation (7.6.4), this implies that

$$dV = 0 = V_w dw + V_r dr_h + V_N dN + V_q dq. \tag{7.6.5}$$

In the standard Rosen-Roback framework with fixed labour supply, using Roy's identity and rearranging the equivalent of the equilibrium condition in equation (7.6.5) gives rise to an analytical expression for the implicit price of quality of life, $p^q = V_q / V_w$, in terms of the observable hedonic gradients, $\frac{dw}{dq}$ and $\frac{dr}{dq}$. However, because labour supply is now a choice variable, the wage rate w no longer corresponds to a pure income term, but now also represents the relative price of leisure l. This requires that Roy's identity is modified appropriately. From the definitions of V_r and V_w, we have

$$-\frac{V_r}{V_w} = -\frac{U_x x_r + U_h h_r + U_l l_r}{U_x x_w + U_h h_w + U_l l_w}. \tag{7.6a}$$

Dividing by U_x and using the first-order conditions $\frac{U_r}{U_x} = h$ and $\frac{U_l}{U_x} = w$ then yields

$$-\frac{V_r}{V_w} = -\frac{x_r + r h_r + w l_r}{x_w + r h_w + w l_w}. \tag{7.6b}$$

Differentiating the budget constraint with respect to r, the numerator reduces to h, and differentiating the denominator with respect to w reduces the denominator to t such that the modified version of Roy's identity can be written as

$$-\frac{V_r}{V_w} = \frac{h}{t}. \tag{7.7}$$

Similarly, it can be shown that $-\frac{V_w}{V_N} = t$ which now permits to re-arrange equation (7.6.5) and find an expression for the implicit price of quality of life in terms of observable hedonic gradients:

$$\begin{aligned} p^q &= -\frac{V_r}{V_w}\frac{dr}{dq} - \frac{V_N}{V_w}\frac{dN}{dq} - \frac{dw}{dq} \\ &= \frac{h}{t}\frac{dr}{dq} - \frac{1}{t}\frac{dN}{dq} - \frac{dw}{dq} \end{aligned} \tag{7.8}$$

Multiplying through by t we get $\tilde{p}^q = t p^q$ which is a measure of the total amount of annual income a worker is willing to pay for an incremental increase in the quality of life:

$$\tilde{p}^q = h\frac{dr}{dq} - \frac{dN}{dq} - t\frac{dw}{dq} \tag{7.9}$$

Empirical estimates for $\frac{dN}{dq}$ can be inferred via some variant of Burtless and Hausman's (1978) standard labour supply equation by expanding $\frac{dt}{dq} = \frac{dN}{dq}\frac{dt}{dN}$ and similarly $\frac{dt}{dq} = \frac{dw}{dq}\frac{dt}{dw}$. From these identities we can then derive the following expression for $\frac{dN}{dq}$ in terms of observable quantities:

$$\frac{dN}{dq} = \frac{\frac{dw}{dq}\frac{dt}{dw}}{\frac{dt}{dN}}, \tag{7.10}$$

where $\frac{dw}{dq}$ is the standard hedonic gradient from wage regressions and $\frac{dt}{dw}$ and $\frac{dt}{dN}$ are directly obtained from estimating a labour supply equation (see section 7.7).

7.7 Empirical issues

This section discusses potential problems with previous empirical estimates of quality of life using microdata and suggests a method for using estimates of the labour supply elasticity to impute an improved estimate of implicit amenity prices.

7.7.1 Hourly wages and the fixed labour supply assumption

Deriving hourly wages from microdata introduces two sources of potential bias. First, there is latent measurement error if hourly wages are constructed from inaccurate measures of labour supply. There are some concerns that data in the category "usual hours worked during the past year", as reported in the 2000 Census, contain errors that create incredible implied hourly wages, a problem that is particularly severe among part-time workers. Baum-Snow and Neal (2009) document concerns that data in the category "usual hours worked during the past year", as reported in the 2000 Census, contain errors that create incredible implied hourly wages for part-time workers, i.e. those who work between one to 34 hours per week. Because of these concerns, we trim our samples and exclude those workers for whom the implied hourly wage is less than $1.50 or greater than $500. We also eliminate workers from the samples who report usually working less that five hours per week

Furthermore, Borjas (1980) highlights the problem of the "division bias" in traditional labour supply estimates due to the appearance of hours on both sides of the equation, i.e $h_i = \alpha + \beta w_i + \varepsilon_i$ where $w_i = y_i / h_i$. The second source of bias is more subtle and arises due to the fact that household labour supply actually varies across locations, whereas the underlying model assumes that each worker supplies a fixed amount of labour to the market. This might be particularly problematic in empirical work that constructs hourly wages using microdata. In this type of work, hourly wages w_i^j for worker i in location j tend to be derived from reported annual wage earnings y_i^j and annual hours worked, i.e. $w_i^j = y_i^j / (k_i^j \times h_i^j)$, where k_i^j are weeks worked per year and h_i^j are hours worked per week. Figure 7.4(a) highlights the substantial variation in the average annual hours worked across different PUMAs. This local variation in labour supply is even larger for male workers compared to female workers. Figure 7.4(b) shows the bias in the hourly wage measure resulting from local variations in the annual hours worked versus applying $(\bar{k} \times \bar{h}) =$ 2,071 hours for the whole sample.

The fact that $\bar{k}^j$ and $\bar{h}^j$ show such large spatial variation compared to the national average introduces a source of wage variation that is inconsistent with the assumptions of the theoretical model (i.e. $\tilde{w}_i^j = y_i^j / (\bar{k} \times \bar{h})$). Consequently, empirical work that constructs hourly wages as w_i^j while assuming $\tilde{w}_i^j$ tends to overstate (understate) the size of the compensation differential in locations where households work more (less) than the national average. This nuance is usually ignored in studies that provide hedonic estimates of quality of life from microdata (e.g. Blomquist, Berger, and Hoehn (1988); Albouy (2008); Chen and Rosenthal (2008)). The resulting consequences for quality-of-life estimates are explored in the next section.

7.7.2 Labour-supply estimates

Most empirical work on labour supply, however, ignores the variation in local prices other than wages. Consequently, estimates of labour supply elasticities are usually derived via some variant of Burtless and Hausman's (1978) hours equation that generates a relationship between log hours and log earnings (or, equivalently, log wages) and non-wage income. Derived directly from Stone-Geary utility functions, this labour-labour supply model has the form:

$$\log h_i = \gamma_0 + \gamma_1 \log w_i + \gamma_2 \log N_i + \boldsymbol{Z}_i' \boldsymbol{\theta} + \nu_i, \tag{7.11}$$

where h_i are hours worked, w_i is the hourly wage, N_i is non-wage income, and $\boldsymbol{Z}$ is a vector of individual attributes and controls, usually including gender, race, education and other variables that affect labour-supply decisions such as non-wage income. Angrist (1991) discusses a number of empirical issues that arise when estimating equation from microdata. Most importantly, w_i needs to be constructed from by dividing annual earnings by annual hours which introduces measurement error.

7.8 Summary and future work

This chapter has outlined several elements of a generalsiation of the Rosen-Roback model of urban location choices with regard to household heterogeneity and labour supply. As a next step, a refinement of the theory section would further derive the relationship between quality of life, labour supply, wages, and rents. Specifically, I have started to work on a model of continuous labor supply where there is an equilibrium condition in which people choose optimal hours of work at which the marginal wage equals the opportunity cost of time. This work would be followed by a briefly review of the data and estimator from chapter 3, repeating the estimation using the adjusted quality-of-life metric and summarising the results.

In an application, future work would look at the reduced form relationship between the generalised adjusted quality-of-life index and local labour supply. This work could take a similar approach to Rosenthal and Strange (2008), and try to find ways to distinguish between their explanations for variation in labour supply, and the new explanations.

In addition to individual labor supply, I think it could be interesting to look at household labor supply. Does the primary earner work more hours in order to afford to live in a high quality-life-area so that the secondary earner can work less and enjoy the amenities offered by a high quality of life? I wonder if we can make testable predictions about the behavior of singles, couples, and families that might parallel the work by Graves and Waldman (1991) on working households vs. retirees.

This work highlights that once one starts thinking about non-market income, one also has to think about leisure. With time spent on generating income (market or non-market), we buy or produce goods and services to meet needs or for simple enjoyment. Time available for leisure obviously affects well-being, and thus the quality of life. Understanding changes in the amount of leisure over time and differences between locations represent an important aspect of understanding social welfare and individual well-being in this context.

References

ALBOUY, D. Y. (2008): "Are Big Cities Really Bad Places to Live? Improving Quality-of-Life Estimates across Cities," Working Paper No. 14472, National Bureau of Economic Research, Cambridge, MA.

ANGRIST, J. D. (1991): "Grouped-data Estimation and Testing in Simple Labor-Supply Models," *Journal of Econometrics*, 47(2), 243–266.

BARTIK, T. J., J. S. BUTLER, AND J.-T. LIU (1992): "Maximum Score Estimates of the Determinants of Residential Mobility: Implications for the Value of Residential Attachment and Neighborhood Amenities," *Journal of Urban Economics*, 32(2), 233–256.

BAUM-SNOW, N., AND D. NEAL (2009): "Mismeasurement of Usual Hours Worked in the Census and ACS," *Economics Letters*, 102(1), 39–41.

BEESON, P. E. (1991): "Amenities and Regional Differences in Returns to Worker Characteristics," *Journal of Urban Economics*, 30(2), 224–241.

BLACK, D. A., N. A. KOLESNIKOVA, AND L. J. TAYLOR (2009a): "Earnings Functions When Wages and Prices Vary by Location," *Journal of Labor Economics*, 27(1), 21–47.

——— (2009b): "Local Price Variation and Labor Supply Behavior," *Federal Reserve Bank of St. Louis Review*, 91(6), 613–626.

BLOMQUIST, G. C., M. C. BERGER, AND J. P. HOEHN (1988): "New Estimates of Quality of Life in Urban Areas," *American Economic Review*, 78(1), 89–107.

BORJAS, G. J. (1980): "The Relationship between Wages and Weekly Hours of Work: The Role of Division Bias," *Journal of Human Resources*, 15(3), 409–423.

BRUECKNER, J. K., J.-F. THISSE, AND Y. ZENOU (1998): "Why is Central Paris Rich and Downtown Detroit Poor?: An Amenity-based Theory," *European Economic Review*, 43(1), 91—107.

BURTLESS, G., AND J. A. HAUSMAN (1978): "The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiment," *Journal of Political Economy*, 86(6), 1103–1130.

CHEN, Y., AND S. S. ROSENTHAL (2008): "Local Amenities and Life-cycle Migration: Do People Move for Jobs or Fun?," *Journal of Urban Economics*, 64(4), 519–537.

EVANS, M. F., AND G. SCHAUR (2010): "A Quantile Estimation Approach to Identify Income and Age Variation in the Value of a Statistical Life," *Journal of Environmental Economics and Management*, forthcoming.

GLAESER, E. L., M. E. KAHN, AND J. RAPPAPORT (2008): "Why Do the Poor Live in Cities? The Role of Public Transportation," *Journal of Urban Economics*, 63(1), 1–24.

GLAESER, E. L., J. KOLKO, AND A. SAIZ (2001): "Consumer City," *Journal of Economic Geography*, 1(1), 27–50.

GRAVES, P. E., AND D. M. WALDMAN (1991): "Multimarket Amenity Compensation and the Behavior of the Elderly," *American Economic Review*, 81(5), 1374–1381.

HECKMAN, J. J., AND T. E. MACURDY (1980): "A Life Cycle Model of Female Labour Supply," *Review of Economic Studies*, 47(1), 47–74.

LEE, S. C. (2010): "Ability Sorting and Consumer City," *Journal of Urban Economics*, 68(1), 20–33.

McMillen, D. P. (2008): "Changes in the Distribution of House Prices Over Time: Structural Characteristics, Neighborhood, or Coefficients?," *Journal of Urban Economics*, 64(4), 573–589.

Nosal, E., and P. Rupert (2007): "How Amenities Affect Job and Wage Choices Over the Life-Cycle," *Review of Economic Dynamics*, 10(3), 424–Ű443.

Redfearn, C. L. (2009): "How Informative are Average Effects? Hedonic Regression and Amenity Capitalization in Complex Urban Housing Markets," *Regional Science and Urban Economics*, 39(2), 297–306.

Roback, J. (1980): "The Value of Local Urban Amenities: Theory and Measurement," Ph.D. thesis, University of Rochester.

——— (1982): "Wages, Rents, and the Quality of Life," *Journal of Political Economy*, 90(6), 1257–1278.

——— (1988): "Wages, Rents, and Amenities: Differences Among Workers and Regions," *Economic Inquiry*, 26(1), 23–41.

Rosen, S. (1979): *Current Issues in Urban Economics* chap. Wage-based Indexes of Urban Quality of Life, pp. 74–104. Johns Hopkins University Press, Baltimore, MD.

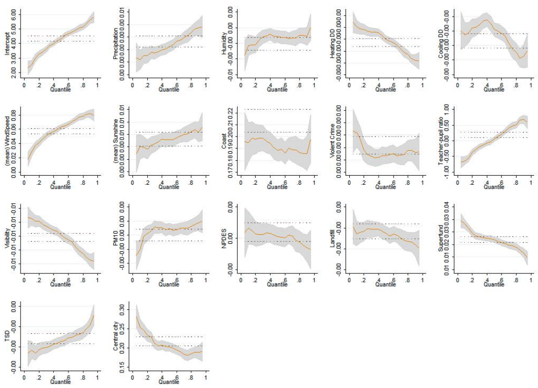
Rosenthal, S. S., and W. C. Strange (2008): "Agglomeration and Hours Worked," *Review of Economics and Statistics*, 90(1), 105–118.

Shan, H. (2010): "Property Taxes and Elderly Mobility," *Journal of Urban Economics*, 67(2), 194–205.

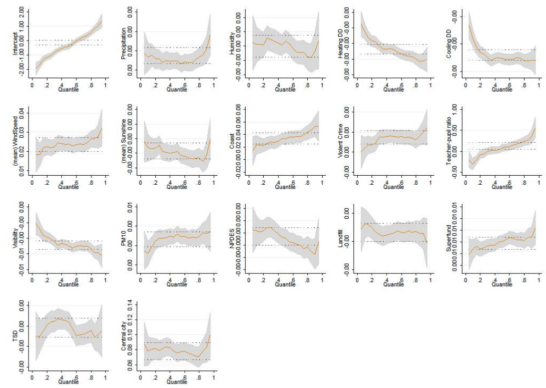
Yang, F. (2009): "Consumption Over the Life Cycle: How Different is Housing?," *Review of Economic Dynamics*, 12(4), 423–443.

Zenou, Y. (2009): *Urban Labor Economics*. Cambridge University Press, New York, NY.

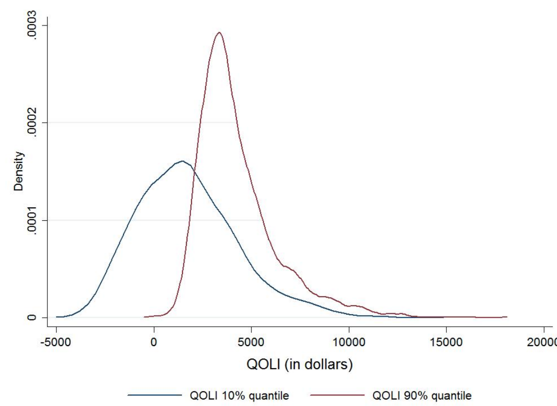
Figure 7.1: Quantile hedonic regressions – metropolitan counties



(a) Housing regression



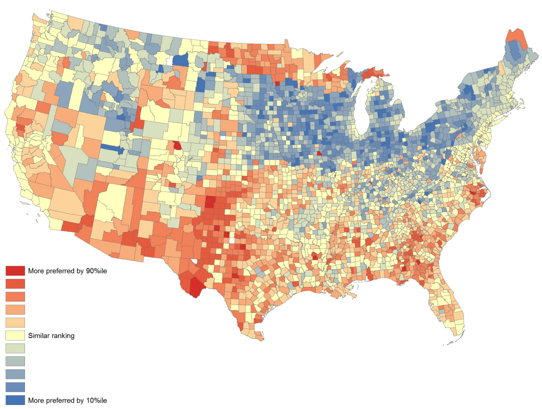
(b) Wage regression



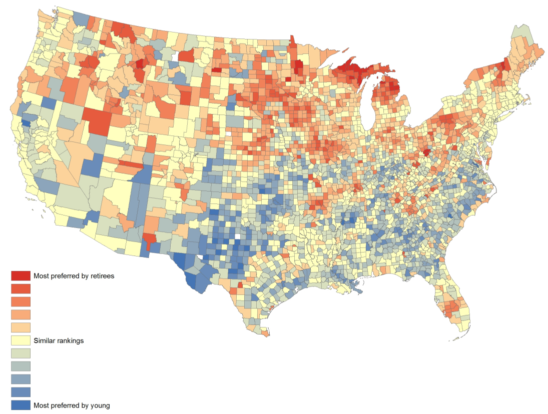
(c) Distribution of QOLI across quantiles

Notes: Variation in the coefficients over the conditional quantiles where the horizontal lines represent the corresponding OLS estimates. The shaded areas around parameter estimates in panels (a) and (b) are 95% confidence intervals. Panel (c) show kernel density estimates of the quality-of-life indices for the highest and lowest quantiles.

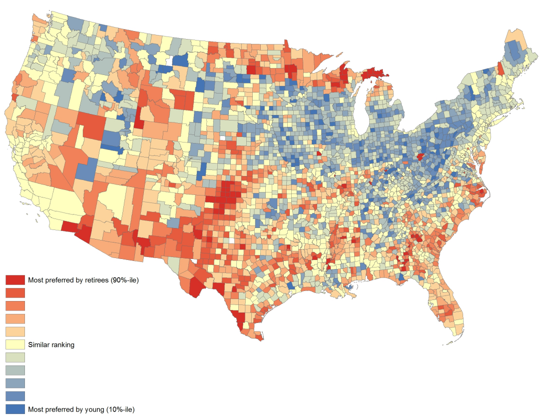
Figure 7.2: Spatial pattern of quality-of-life preferences across age and housing consumption



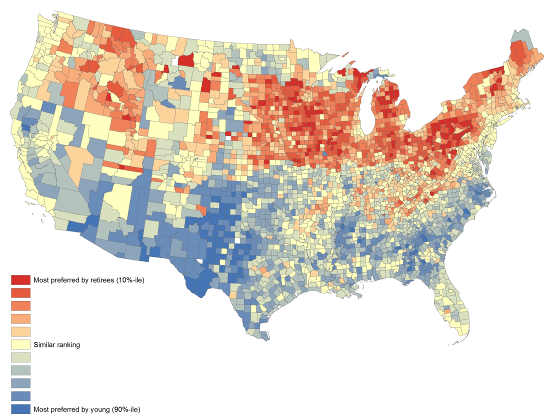
(a) All ages



(b) Retirees and young



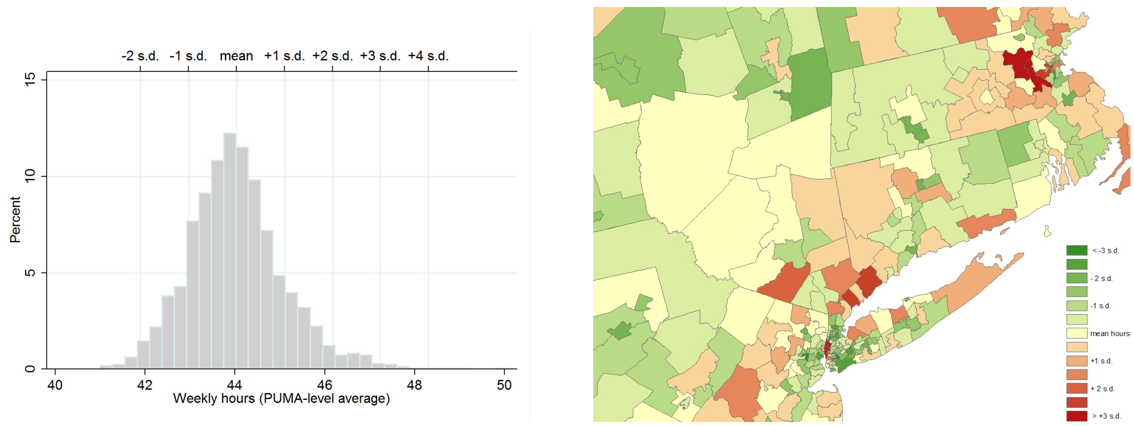
(c) Preferences of 90%ile retirees and 10%ile young



(d) Preferences of 10%ile retirees and 90%ile young

Source: Authors' calculations.

Figure 7.3: Local variation in the labour supply behaviour of full-time workers

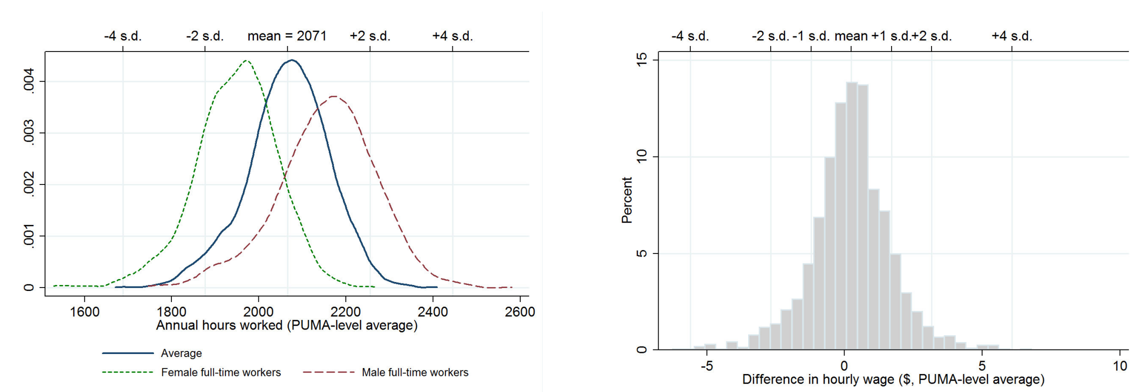


(a) Differences in weekly hours worked

(b) Spatial variation of labour supply

Notes: Full-time workers are defined using the standard BLS definition as persons who work 35 hours or more per week. *Source*: Author's calculations from PUMS data.

Figure 7.4: Imputed hourly wages



(a) Local variation of full-time workers' annual hours worked

(b) Bias in hourly wage measure

Notes: Full-time workers are defined using the standard BLS definition as persons who work 35 hours or more per week. Hourly wages w_{ij} for full-time worker i in location j are derived from reported annual wage earnings y_{ij} and annual hours worked, i.e. $w_{ij} = y_{ij}/(k_{ij} \times h_{ij})$, where k_{ij} are weeks worked per year and h_{ij} are hours worked per week. Figure 7.4(a) illustrates the variation of annual hours worked across gender. Figure 7.4(b) shows the bias in the hourly wage measure due to local variations in the annual hours worked versus applying $(\bar{k} \times \bar{h}) = 2{,}071$ hours for the whole sample. Hotelling's T^2 generalised test of a zero mean is rejected at the 1% level. *Source*: Author's calculations from PUMS data.

Chapter 8

Making quality of life endogenous

8.1 Introduction

The main purpose of this chapter is to outline several aspects of future work where the quality-of-life metrics presented in chapters 3 and 7 are integrated into the larger context of the regional economy. This would imply that quality-of-life would become an endogenous quantity with respect to a number of macroeconomic processes that determine local economic outcomes. In terms of the graphical model of regional linkages shown in figure 1.1 of chapter 1, the aim of making quality of life endogenous to the system could perhaps – somewhat simplistically – be described as having the blue arrows that emanate from the "amenities" in the upper portion of the diagram run in both directions. Different aspects of public policy are particularly relevant examples of instances where the quality of life simultaneously is both affected by and is itself influencing economic activity. From optimal city size and housing supply restrictions due to land-use regulation to the regional impacts of housing-related income tax benefits and environmental policy, quality of life interacts with the decision making processes of both economic agents and policy makers.

This chapter briefly discusses some preliminary evidence of how the quality-of-life metrics presented in the previous chapters relate to several important policy questions in urban economics and regional science. Section 8.2 examines the relationship between urban size and the resulting quality of life, briefly revisiting the discussion of optimal agglomeration and urban scale. Section 8.3 provides some preliminary graphical evidence on the possible connections between regional divergence in the quality of life and specific federal, state and local policies. Section 8.4 considers the nexus between financial markets and regional goods markets where risk premia in the housing market might act as a conduit for local differentiation in terms of non-market goods such as the quality of life.

8.2 Optimal urban size

Well into the early 1970s, many urban scholars across disciplines shared a somewhat Dickensian view of large cities. According to this point of view, many of the negative externalities of economic growth were connected with urbanisation and congestion and the occupational and residential transformation of the Industrial Revolution is viewed as a necessary evil to have enjoyed the fruits of technological progress. James Tobin and William Nordhaus, for example, suggest that "[...] some portion of the higher earnings of urban residents may simply be compensation for the disamenities of urban life" (Nordhaus and Tobin, 1972, p.552).

On the other side of the debate, one of the fathers of modern urban economics, William Alonso, famously pleaded to "[...] slow the rush to public acceptance of disamenities [...] as the defenders of the great cities we take Faustian positions; we see the positive values of civilisation flowing from them" (Alonso, 1976, p.51). At the same time, some economic historians increasingly argued that William Blake's *Dark Satanic Mills* view of the Industrial Revolution – that the quality of urban life as it influenced the common man was deteriorating following 1750 – simply did not wash. Or as Jeffrey Williamson asks: "What did the labourer forego in leaving some rural 'Sweet Auburn' for some ugly, urban Sheffield during the Industrial Revolution?" (Williamson, 1981, p.77). The question of the relationship between urban size and quality of life remains an active one today.

While chapter 3 has provided fresh empirical insights on the intra-urban variation in quality-of-life, we have also seen that some urban heterogeneity may depend on city size. Most previous work, however, has neglected scale-related intra-urban diversity by assigning the same observational weight to a one-county metro like Pueblo, Colorado and to a global mega-city like New York City which combines the extremes of suburban affluence (Hunterdon county, New Jersey and Rockland county, New York) and innercity poverty (Bronx County, NY). Even for large metropolitan areas of similar size such as Atlanta, Boston, or Miami, metro-level aggregates of quality-of-life measures are likely to mask significant intra-urban differences that are influenced by urban structure and population densities.

8.2.1 Intraurban variation in the quality of life

Grouping cities by population size, metropolitan areas with the largest intraurban variation in the quality-of-life at each level of urban scale. Table 8.1 lists the five metropolitan areas with the largest intraurban variation in the quality-of-life at each level of urban scale. The MSA-level aggregate quality-of-life values are calculated as population-weighted averages of the relevant county quality-of-life premia. A set of separate quality-of-life rankings are obtained for the 358 MSAs on the basis of these population-weighted average QOLIs.[1]

Table 8.1: Intra-urban variation of quality-of-life indices

Metropolitan Statistical Area Top/bottom county	Population ('000s)	Range of QOLI*	QOLI†	Rank‡
Very large metro areas (population > 3 million)				
Atlanta-Sandy Springs-Marietta, GA	4,248	$38,540	$11,930	*101*
Fulton County, GA			$38,540	107
Dawson County, GA			$0	1,085
Washington-Arlington-Alexandria, DC-VA-MD-WV	4,796	19,689	14,387	*36*
Montgomery County, MD			41,587	32
Jefferson County, WV			21,898	1,081
New York-Northern New Jersey-Long Island, NY-NJ-PA	18,300	11,828	18,309	*10*
Nassau County, NY			47,310	5
Pike County, PA			35,482	304
Chicago-Naperville-Joliet, IL-IN-WI	9,098	9,308	14,296	*39*
Lake County, IL			42,527	25
Jasper County, IN			33,219	655
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	5,687	7,734	14,152	*41*
Montgomery County, PA			41,424	35
Salem County, NJ			33,691	573
Large metro areas (population > 1.5 million)				
Virginia Beach-Norfolk-Newport News, VA-NC	1,576	$14,526	$11,840	*107*
Williamsburg city, VA			$42,644	23
Currituck County, NC			$28,118	1,067
St. Louis, MO-IL	2,699	11,156	11,006	*151*
St. Louis County, MO			36,905	182
Calhoun County, IL			25,749	1,074
San Jose-Sunnyvale-Santa Clara, CA	1,736	9,618	24,878	*1*

continued on the next page

[1] Interestingly, spatial aggregation matters for the quality-of-life indices. Our rankings have a minimal rank correlation (0.09-0.15) with the original BBH rankings when measured at the county level. However, when we aggregate rankings to the MSA-level, the rank correlation increases to 0.23. This further highlights the importance of intra-urban differences in the quality-of-life which are not captured by studies that use MSA-level aggregates

Metropolitan Statistical Area Top/bottom county	Population ('000s)	Range of QOLI*	QOLI†	Rank‡
Santa Clara County, CA			49,032	4
San Benito County, CA			39,414	86
Kansas City, MO-KS	1,836	8,218	12,033	*95*
Johnson County, KS			37,817	135
Bates County, MO			29,599	1,048
Denver-Aurora, CO	2,158	7,826	14,263	*40*
Gilpin County, CO			41,446	34
Clear Creek County, CO			33,620	592
Mid-sized metro areas (population > 0.5 million)				
Richmond, VA	1,097	$10,877	$8,594	*296*
Goochland County, VA			$37,095	171
Colonial Heights city, VA			$26,218	1,072
Rochester, NY	1,038	6,605	13,995	*44*
Monroe County, NY			39,075	94
Livingston County, NY			32,470	785
Syracuse, NY	6,50	6,476	11,926	*102*
Onondaga County, NY			37,542	149
Oswego County, NY			31,066	965
Oklahoma City, OK	1,095	6,433	9,184	*260*
Oklahoma County, OK			34,516	440
Logan County, OK			28,083	1,068
Hartford-West Hartford-East Hartford, CT	1,149	6,107	13,402	*54*
Middlesex County, CT			41,372	37
Tolland County, CT			35,265	335
Small metro areas (population < 0.5 million)				
Reno-Sparks, NV	343	$24,517	$18,406	*9*
Washoe County, NV			$42,508	26
Storey County, NV			$17,991	1082
Cleveland, TN	104	22,891	6,077	*354*
Bradley County, TN			33,469	616
Polk County, TN			10,578	1,084
Victoria, TX	112	16,441	7,908	*322*
Calhoun County, TX			32,923	715
Goliad County, TX			16,481	1,083
Glens Falls, NY	124	10217	11,354	*128*
Warren County, NY			40,229	58
Washington County, NY			30,012	1,032
Beaumont-Port Arthur, TX	385	8,565	7,564	*328*
Hardin County, TX			38,434	110
Orange County, TX			29,869	1,039

Notes: * The quality-of-life index values are based on the improved specifications labelled "$QOLI_5$" in the last column of table 3.8. The QOLI range is the difference in implicit expenditure between the most desirable and the least desirable county within a given MSA. QOLI values are rescaled to $0 for lowest ranking county. † The QOLI values for the MSA are the population-weighted averages of the relevant county QOLI values. ‡ Separate QOLI rankings are obtained for the 358 MSAs on the basis of the population-weighted average QOLI values. The complete list of intra-urban quality-of-life differentials for all MSAs is available in the online appendix.

For the group of the largest US cities with a total population in excess of three million peoples, Atlanta, GA metropolitan area has the largest differential in terms of the quality-of-life premia between the most amenable and least amenable county. Separated by a quality-of-life gap of $38,540, 107th-ranked Fulton County, GA which contains Atlanta is one of the core counties of this large metro, whereas 1,085th-ranked Dawson County, GA is at the outer-fringe of this megaregion that covers 8,376 sq. mi., an area almost the size of New Jersey. The Washington, DC metro area shows similar intra-urban quality-of-life disparities: although less than 50 miles apart, living in 32th-ranked mostly suburban Montgomery county in Maryland as opposed to 1,081st-ranked Jefferson county in West Virginia is worth $19,689 to the average household.

By contrast, the gradient of intraurban quality-of-life differentiation in the nation's second most desirable MSA, San Francisco which contains first-ranked Marin County, is considerably lower at $6,572. These large differences notwithstanding, the relationship between intraurban heterogeneity does not appear to depend on urban size in a significant manner. Bar a few exceptions, table 8.1 shows that across all groups of urban scale the most substantial intraurban differences in the quality-of-life are bounded between $2,500 and $10,000. The complete list of intra-urban quality-of-life differentials for all MSAs is available in the online appendix.

8.2.2 Quality of life and urban scale

While metropolitan disparities thus seem largely unrelated to scale, the aggregate MSA quality-of-life indices reveal some positive correlation with urban size, which is broadly consistent with Herzog and Schlottmann (1993) or Albouy (2008). The relationship between urban scale and quality of life has a long-standing tradition in the urban literature where the discussion is linked to the notion of identifying a 'optimal'-sized urban area (cf. Richardson, 1972, 1973). From an urban economic perspective, larger metropolitan areas exist because of productivity advantages associated with the concentration of economic activity cities. At the same time, increasing population density and large urban populations create diseconomies and negative externalities due to congestion and (environmental) degradation, creating size-related costs of urbanisation. In this framework, optimal city size is reached when both the benefits of size-related amenities and productivity gains and the cost from disamenities are equally balanced. Although the augmented Roback-HBB framework incorporates such agglomeration dynamics via productivity effects, the model makes no a priori predictions about the relationship between quality of life and urban size which in the end remains an empirical matter.

In contrast to the intuitive appeal of popular rankings, most previously published hedonic quality-of-life indices do not vary with city size or they tend to have a negative relationship.[2] This is documented in table 3.9, listing the most prominent papers that have produced theoretically consistent rankings for which the relationship between quality-of-life and population size is then computed. Table 3.9 also reports how these rankings correlate with our own new indices and add the the rankings from the *Places-Rated Almanac* and Sperling's *Cities Ranked & Rated* as a basis for comparison.

In the original BBH rankings, only six counties belonged to a metropolitan area with more than 1.5 million people. Thirteen of the top 25 counties were part of a city with less than half a million inhabitants. In our new rankings, however, small is no longer necessarily beautiful as this pattern is completely reversed: Seventeen counties are part of metropolitan areas with at least three million people and only three counties belong to small cities where the population is less than half a million. The positive relationship between city size and our new quality-of-life rankings is illustrated in figure 8.1.

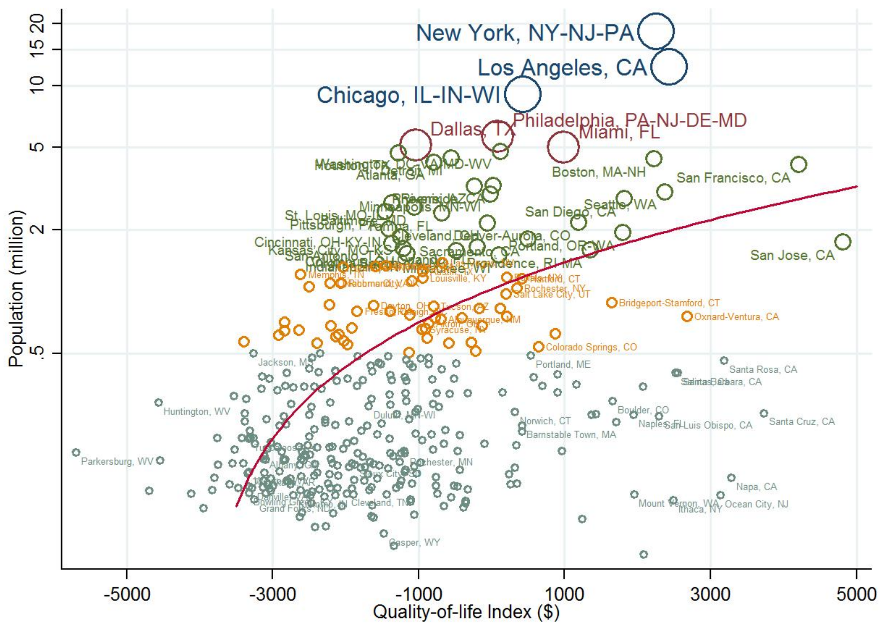
As big cities are thus not inevitably synonymous with the pessimistic image of Gotham City (or Sheffield), our results also point towards an increasing importance of consumption amenities which implies that the willingness to pay for quality of life increases with population density (e.g. Glaeser, Kolko, and Saiz, 2001; Rappaport, 2008).

This qualitative difference in the amenity-city size relationship between our new rankings and those in BBH can also be illustrated graphically, using an intuitive test introduced by Albouy (2008). Figure 8.2 shows the regression line of the housing-expenditure differentials predicted by wage differentials. Recall from chapter 3 that equation (A-6) defines the condition for locational equilibrium, whereby – for a given level of quality of life – increases in wages are precisely offset by higher cost of living, such that real income remains constant.

At the average level of quality of life, the locus of the locational equilibrium combinations of housing expenditure and compensating differentials goes through the origin. The precise slope of this line depends on the relative weight of wages relative of housing expenditure implied by equation (A-6) and I plot the locus that corresponds to our parameterisation (green long-dashed line) and the steeper locus compatible with BBH (orange short-dashed line) in Figure 8.2. Since more (less) amenable locations are positioned above (below) this locus, the vertical distance between the regression line and the equilibrium locus produces the empirical relationship between nominal wages and quality of life. Similarly, the more small (large) cities lie below (above) this line, the stronger the positive association between quality of life

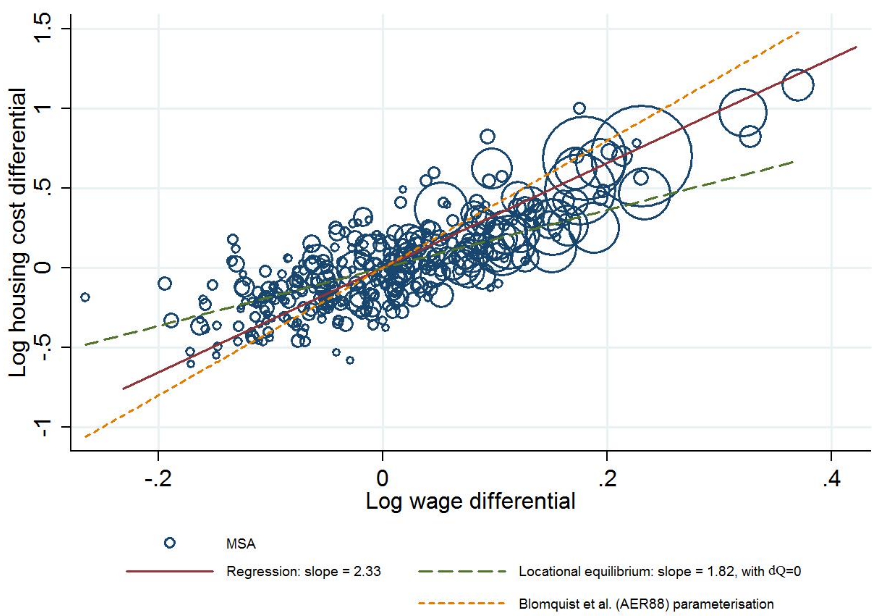
[2] Stover and Leven (1992) and Albouy (2008) are the exceptions to this rule.

Figure 8.1: Quality of life and urban size



Notes: The quality-of-life values for the MSAs are the population-weighted averages of the relevant county quality-of-life values. The size of the markers is proportional to a given MSA's urban population, where blue indicates the largest MSAs with a population in excess of 9 million inhabitants, red are very large metro areas with more than 5 million in habitants, green are mid-sized MSAs with more than 1.5 million inhabitants, orange are smaller MSAs with a population of at least 0.5 million people and gray indicates the smallest metropolitan areas with a population of less than 500,000 inhabitants. The red line indicates a quadratic fit for urban size and quality of life.

Figure 8.2: Quality of life and locational equilibrium



Notes: Markers represent individual MSAs with the size proportional to the total MSA population. The condition for the locational equilibrium is given by equation (A-6) which defines the locus along which increases in wages are precisely offset by higher cost of living, such that real income remains constant for a given level of quality of life. The dashed lines through the origin represent this locus for different parameterisations at the average level of quality of life. More (less) amenable locations are positioned above (below) the locus. The vertical distance between the regression line and the equilibrium locus produces the relationship between nominal wages and quality of life.

and urban scale. The steeper slope of the BBH locus means that the bulk of small MSAs lie above the locus and most larger cities are located below, hence the “small is beautiful effect”. The opposite is the case with regard to the equilibrium locus that reflects our calibration, indicating that larger cities tend to be more desirable places to live.

8.3 Quality of life and public policy

This section outlines some preliminary graphical evidence on the relationship between quality of life and different levels of public policy. At the federal level, I investigate the correlation between local differences in the burden of fiscal policy and quality of life. At the state and local level, amenity-driven quality-of-life differentials are linked to environmental policy and land-use regulation.

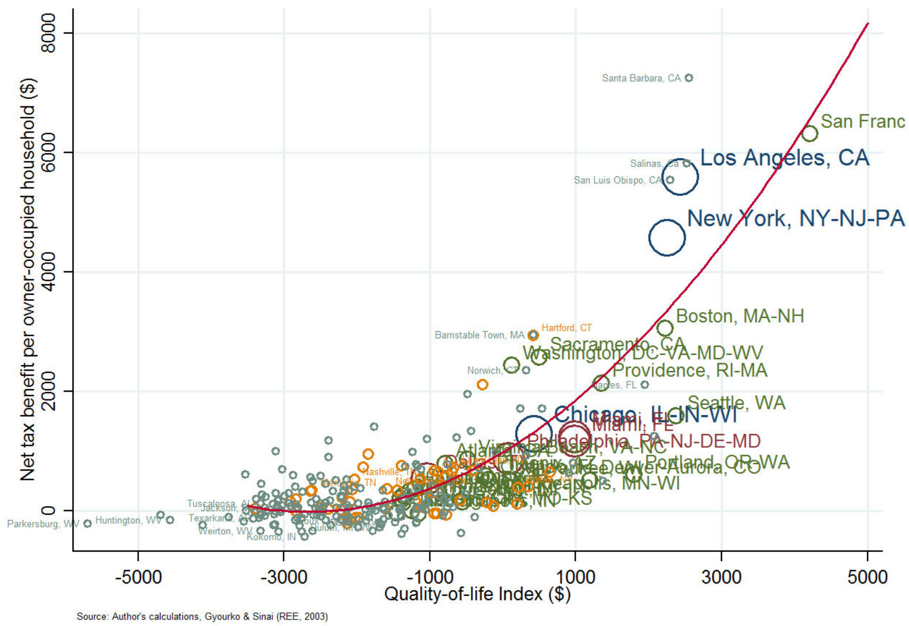
8.3.1 Quality of life and housing-related income tax benefits

With some 40 million American households each claiming an average of $9,500 in mortgage interest deductions (MID) and almost $3,000 in property tax deductions in 2003, the subsidy of homeownership is indeed one of the most prominent features of the American tax code altogether. Gyourko and Sinai (2003) calculated how tax subsidies to owner-occupied housing are distributed spatially across the United States and find a remarkable skewness with regard to the geographical the burden of this policy. Most interestingly, perhaps, is the their result that the largest urban areas tend to be the biggest net beneficiaries of this federal policy.

The net income tax benefits due to homeownership are is plotted on the vertical axis in figure 8.3 and on the horizontal axis I am plotting the population-weighted quality-of-life indices for each MSA. Gyourko and Sinai (2003) assume that the program is self-funded on a lump sum basis. They assess a lump-sum cost per household equal to the $2,092 national average benefit. The net program benefits then correspond to the difference between that amount and the effective gross benefits received. In a few, mostly smaller, remote MSAs, households actually receive less benefits that the lump-sum cost of financing; i.e. they would be better off if mortgage interest, benefits from owner-occupied rents and property tax deductions were not allowed. Consequently, the net income tax benefits for a few small MSAs, such as Parkersburg, West Virginia and Huntington, West Virginia, are actually negative.

In addition to these spatial inequality of the distributional effect of housing-related federal tax benefits, figure 8.3 also shows a strong positive link between the size of the net benefit and quality of life.

Figure 8.3: Quality of life and housing-related income tax benefits



Notes: Net income tax benefits due to homeownership are taken from Gyourko and Sinai (2003) who assume that the program is self-funded on a lump sum basis ($2,092 national average benefit per household). Quality-of-life indices for each MSA are derived from the corresponding population-weighted county indices.

The larger the per household average benefit, the higher the quality of life in the respective urban area. Gyourko and Sinai (2003) calculate that California receives $15 billion in net benefits annually from the rest of the country. At the metropolitan area level, the top five areas of Los Angeles-Riverside-Orange County, New York-Northern New Jersey, San Francisco-Oakland-San Jose, Boston-Worcester-Lawrence and Washington, D.C-Baltimore received the vast majority of all positive net inflows at the MSA level ($22.7 billion out of $26 billion).

Coincidence of these large regional transfer payments in cities with the highest levels in quality of life might indeed suggest that federal housing housing policy is actually subsidising the quality of life where it is already above average. Beyond the large spatial disparities with regard to their geographic incident, it might also be argued that these subsidies reflect the provision of "hidden welfare" that sustain a high quality of life for the urban wealthy through substantial tax benefits on large homes (Chaves, Knox, and Bieri, forthcoming). A detailed discussion of the lack of adjustment in the federal income tax code for differences in regional living costs is contained in Knoll and Griffith (2003) and Albouy (2009).

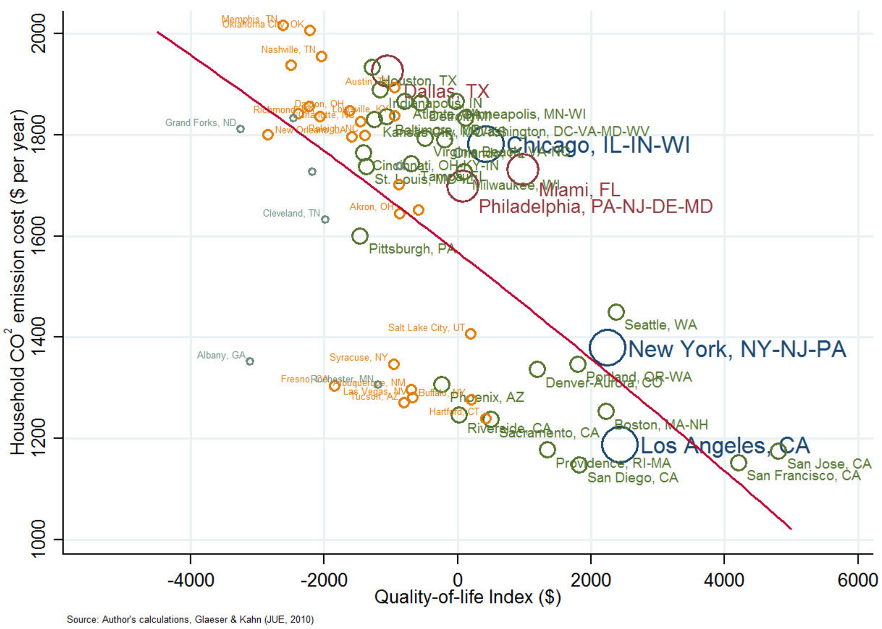
Despite some remote prospects of an overhaul of the system of housing-related tax benefits in the US, specific aspects of the government's policy responses to the recent financial crisis provide further evidence of an implicit federal subsidy for homeowners in areas with high quality of life. In February 2010, President Obama announced a plan to direct $1.5 billion in taxpayer money to five state housing finance agencies to help them develop new programs for addressing the housing crisis in their communities, according to a senior administration official. The five states – California, Nevada, Arizona, Michigan and Florida – have been among the hardest hit by the housing crisis and have seen home values decline more than 20 percent. The initiative will be financed through the government's Troubled Assets Relief Program (TARP).

8.3.2 Quality of life and environmental policy

Carbon dioxide emissions may create significant negative externalities because of global warming. At the same time, however, CO_2 abatement policies vary significantly across the United States. Glaeser and Kahn (2010) quantify the carbon dioxide emissions associated with new construction in different locations across different metropolitan areas.

The annual cost of emissions per household from driving, public transit, home heating, and electricity usage is plotted on the vertical axis in figure 8.4 and – as before – metropolitan quality of life is shown on the horizontal axis. While the lowest emissions areas are generally in California and that the

Figure 8.4: Quality of life and the greenness of cities



Notes: Quality-of-life indices for each MSA are derived from the corresponding population-weighted county indices.

highest emissions areas are in Texas and Oklahoma, this graph also shows that the presence of emissions (greenness of cities) has a strong negative (positive) correlation with urban quality of life. At the same time, however, this relationship appears to hold across cities of all sizes, clearly emphasising the direct link between progressive environmental policy and locational desirability.

8.3.3 Quality of life, the housing-supply restriction and land-use regulation

From a policy perspective, open space represents one of the primary environmental concerns in urban areas. Land-use regulation thus has a direct influence on the quality of life in urban areas by the way in which it influences the shaping of the built environment. The restriction of housing supply due to zoning regulation is perhaps one of the most significant channels through which local governments have a direct effect on the forces of urban agglomeration, especially in a country like the United States with significant local variation in land use and other regulatory practices. Green, Malpezzi, and Mayo (2005) find strong evidence that differences in supply elasticities indeed stem mainly from differences in urban form and urban land-use regulation. The related issues of growth management and the provision of landscape amenities also play a central role in the policy mix of local governments and have important consequences for regional disparities in the quality of life.[3]

The Wharton Residential Land Use Regulation Index (WRLURI) developed by Gyourko, Saiz, and Summers (2008) provides an aggregate measure of regulatory constraint on development that allows to rank areas by the degree of control over the residential land use environment. Figure 8.5 plots the relationship between the strength of land-use regulation and the quality of life in urban areas. The substantial amount of positive covariation between the two variables suggests that locations with a more restrictive land-use regime are also more desirable from the perspective of the quality of life. Again, this relationship does not appear to be significantly dependent on urban size, although the most restrictively regulated localities are smaller towns in New England where home rule specific-aspects of land-use regulation afford municipalities a substantial amount of self determination.

While cities in the West traditionally tend to be the least regulated, anti-sprawl land-use regulation in California (e.g. the state's Senate Bill 375) and the distinct policy innovations of the Pacific Northwest (e.g. Portland's urban growth boundary) tend to coincide with a higher quality of life.[4] Because of the popular

[3] See Walsh (2007) for a discussion of the equilibrium impact of open space protection and growth control policies on the entire metropolitan landscape.

[4] California Senate Bill 375 is typically referred to as the Anti-Sprawl Bill and requires planners and policymakers to develop

Figure 8.5: Quality of life and land-use regulation

Notes: The Wharton Residential Land Use Regulation Index (WRLURI) provides an aggregate measure of regulatory constraint on development that allows to rank areas by the degree of control over the residential land use environment (see Gyourko, Saiz, and Summers, 2008). WRLURI for each MSA were constucted from town-specific indices. Quality-of-life indices for each MSA are derived from the corresponding population-weighted county indices.

notion that more restrictive land-use regulation might lead to higher house prices, there could be the worry that quality-of-life metrics are artificially inflated in areas with tighter regulation. However, recent evidence by Glaeser and Ward (2009) dispels such concerns showing that more stringent regulations are only associated with higher prices when contemporary density and demographics are not controlled for, but not these contemporaneous controls are added.

8.4 Real estate market equilibrium, housing risk premia and quality of life

This final section considers the nexus between financial markets and regional goods markets where risk premia in the housing market might act as a conduit for local differentiation in terms of non-market goods such as the quality of life. Like the previous sections in this chapter, it is intended as being purely exploratory in nature, attempting to join two previously unrelated strands of literature in regional science, namely real estate finance and urban economics. As the case studies in Hall's (1982) "Great Planning Disasters" so vividly illustrate, real estate markets are especially well-suited to illustrating the complex linkages and interdependencies between the planning process, local public finance, and agents across different local, national and even international markets.

8.4.1 Real estate price formation, expectations and planning

I begin by recalling that figure 1.2 in the introductory chapter provides a schematic overview of how the monetary transmission mechanism operates at the level of the regional economy. In this stylised model, equilibrium real estate prices P_t provide the key linkage between the relative price of money, i.e. the local interest rate i, and the relative price of housing services, i.e. rent r:

$$P_t = \mathsf{E}\left[\sum_{t=1}^{T} \frac{r_t}{(1+i_t)^t} + \frac{P_T}{(1+i_t)^T}\right] - \sum_{t=1}^{T} \frac{\psi_t}{(1+i_t)^t} - \frac{\Psi_T}{(1+i_t)^T}. \tag{8.1}$$

Equation (8.1) highlights that house prices are not simply a function of spot rents r_t, but that they also

meaningful solutions to curb sprawl, reduce driving, and promote growth in areas that will have the least impact on the environment.

depend on the expectations of market participants regarding the capitalisation of future rents, $\mathsf{E}[\cdot]$, and uncertainty premia ψ_T and Ψ_T. This opens up a central role for urban and regional planning as anchors for expectations. Moreover, the planning process does not only directly influence real estate prices by the extent to which it helps to co-determine the expectations of market participants. Planning also affects real estate prices by lowering the overall level of uncertainty in the real estate market. Furthermore, equation (8.1) also shows how different investment horizons might drive a wedge between the break-even rates of large-scale land developers operating in national markets on the one hand and a locally-governed social welfare maximiser one the other hand.

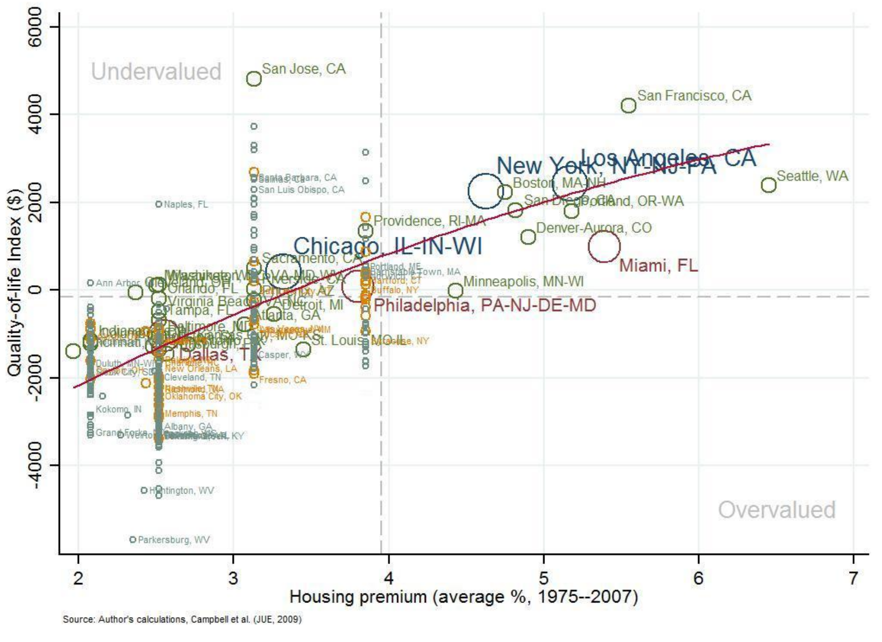
While work by Meade (1970) and Miller (1980) recognises that planning improves overall economic decision making by reducing uncertainty, the role of planning in forming expectations in the real estate market has not been studied in the urban literature. Given that the process of public decision at all levels of government is ultimately characterised by either policy rules or policy discretion, there appear to be a number of unexplored parallels to the vast theoretical literature on the effects of the time inconsistency of optimal plans, a literature that starts with Kydland and Prescott (1977) and has subsequently transformed the theoretical framework of monetary policy analysis.

8.4.2 Monetary transmission mechanism and real estate

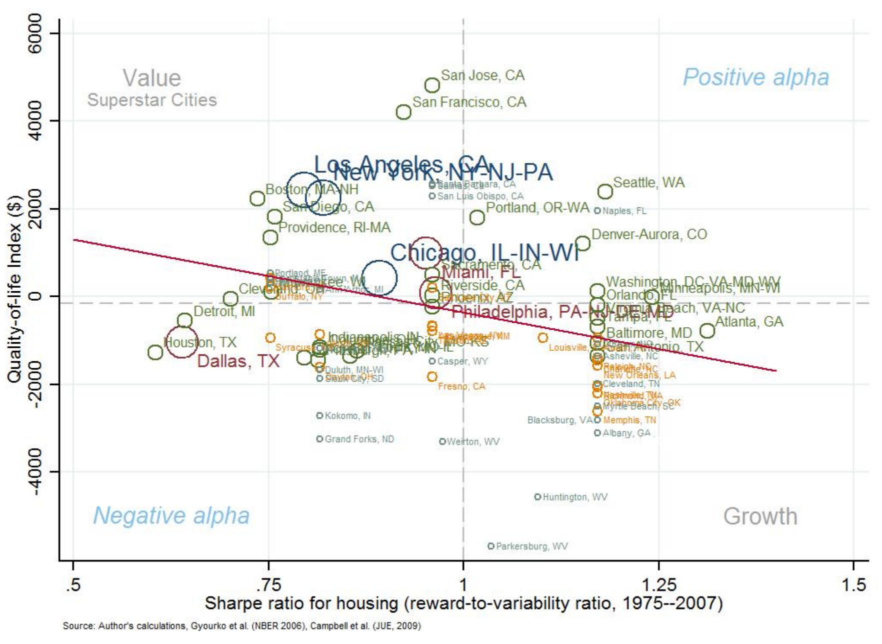
Real estate markets thus "open up the window" to the monetary sector as housing markets are tightly connected to the business cycle (see e.g. Ghent and Owyang (2010) on the relative importance of national vs. regional drivers in the real estate markets). Furthermore, housing markets are also the primary conduit through which the regional effects of monetary policy are transmitted (Fratantoni and Schuh, 2003). In this context, national policy can have very different local effects, for example such as the reduction in inflation which can trigger run-ups in regional housing prices if people suffer from money illusion. Brunnermeier and Julliard (2008) illustrate that investors who decide whether to rent or buy a house by simply comparing monthly rent and mortgage payments do not take into account the fact that inflation lowers future real mortgage costs.

Figures 8.6(a) and 8.6(b) visualise how some of these monetary phenomena might be linked to regional goods markets via tradable goods, such as real estate, and non-tradables, such as the amenity-driven quality of life. In this context, it is useful to decompose real estate returns φ into three separate components, namely risk-free returns i, a housing premium π and rent growth Δr, and where $\varphi = i + \pi$, as originally proposed by Campbell and Shiller (1988):

Figure 8.6: Returns to housing, risk premia and quality of life



(a) Nominal housing premia and regional variation in the quality of life



(b) Risk-adjusted housing risk and regional variation in the quality of life

Notes: The housing premium π are defined as the excess returns to housing φ over and above the returns on a risk-free asset i, such that $\varphi = i + \pi$. The Sharpe Ratio measures the excess return per unit of risk, i.e. $S = \phi/\sigma$, where σ is the standard deviation of the housing premium. *Sources*: Author's calculations using average housing premia for MSAs from Campbell, Davis, Gallin, and Martin (2009).

$$
\begin{aligned}
\log\left(\frac{r_t}{P_t}\right) &= \kappa + \mathrm{E}\left[\sum_{j=0}^{\infty}\rho^j \varphi_{t+1+j} - \sum_{j=0}^{\infty}\rho^j \Delta r_{t+1+j}\right] \\
&= \kappa + \underbrace{\mathrm{E}\sum_{j=0}^{\infty}\rho^j i_{t+1+j}}_{\text{risk-free returns}} + \underbrace{\mathrm{E}\sum_{j=0}^{\infty}\rho^j \pi_{t+1+j}}_{\text{housing premium}} - \underbrace{\mathrm{E}\sum_{j=0}^{\infty}\rho^j \Delta r_{t+1+j}}_{\text{rent growth}} \qquad (8.2)
\end{aligned}
$$

In figure 8.6(a), I plot the quality of life against the nominal housing premia for metropolitan areas π which are obtained from a recent study by Campbell, Davis, Gallin, and Martin (2009). The positive correlation between housing premia and quality of life suggests that homeowners in more desirable areas are systematically rewarded more than those who live in cities with a lower quality of life. In line with fair-value analysis, a technique that is commonly applied in finance, this relationship would also suggest that the housing market is undervalued (overvalued) in cities that have a significantly above-average (below-average) quality of life and a below-average (above-average) housing premium.

However, looking at absolute nominal risk premia alone might be misleading; figure 8.6(b) demonstrates that the relationship between quality of life and systematic rewards in the housing market turns negative, once we adjust for the volatility of those returns to housing. Indeed, once differences in the variability of the local housing market are accounted for, quality of life and the risk-adjusted local housing premium appear to be substitutes rather than complements. This is an interesting preliminary result, since it suggests that the superior risk-adjusted returns on investments in the housing market in less desirable locations actually compensate households for the relatively lower quality of life in these cities.

Because real estate is the quintessential durable good – it can be built quickly, but disappears slowly – urban decline is not the mirror image of growth. Indeed, these asymmetries in urban development imply that "too big to fail" also applies to real estate markets. At the same time, however, the role of the "lender of last resort" is rarely well-defined.

The fragments presented in this chapter outline interesting, previously unexplored aspects of the extent to which the political economy of public decision making has the potential of being (mis)aligned with commercial interests or private preferences. It is to the synthesis of these elements that I wish to contribute in my future academic work.

8.A References

ALBOUY, D. Y. (2008): "Are Big Cities Really Bad Places to Live? Improving Quality-of-Life Estimates across Cities," Working Paper No. 14472, National Bureau of Economic Research, Cambridge, MA.

——— (2009): "The Unequal Geographic Burden of Federal Taxation," *Journal of Political Economy*, 117(4), 635–667.

ALONSO, W. (1976): "Urban Disamenities," *Society*, 13(4), 51–53.

BRUNNERMEIER, M. K., AND C. JULLIARD (2008): "Money Illusion and Housing Frenzies," *Review of Financial Studies*, 21(1), 135–180.

CAMPBELL, J. Y., AND R. J. SHILLER (1988): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1(3), 195–228.

CAMPBELL, S. D., M. A. DAVIS, J. GALLIN, AND R. F. MARTIN (2009): "What Moves Housing Markets: A Variance Decomposition of the Rent-Price Ratio," *Journal of Urban Economics*, 66(2), 90–102.

CHAVES, E., P. L. KNOX, AND D. S. BIERI (forthcoming): *International Perspectives on Suburbanization: A Post-Suburban World* chap. The Restless Landscape of Metroburbia, pp. 42–76. Routledge, London.

FRATANTONI, M., AND S. SCHUH (2003): "Monetary Policy, Housing, and Heterogeneous Regional Markets," *Journal of Money, Credit and Banking*, 35(4), 557–589.

GHENT, A. C., AND M. T. OWYANG (2010): "Is Housing the Business Cycle? Evidence from US Cities," *Journal of Urban Economics*, 67(3), 336–351.

GLAESER, E. L., AND M. E. KAHN (2010): "The Greenness of Cities: Carbon Dioxide Emissions and Urban Development," *Journal of Urban Economics*, 67(3), 404–418.

GLAESER, E. L., J. KOLKO, AND A. SAIZ (2001): "Consumer City," *Journal of Economic Geography*, 1(1), 27–50.

GLAESER, E. L., AND WARD (2009): "The Causes and Consequences of Land Use Regulation: Evidence from Greater Boston," *Journal of Urban Economics*, 65(3), 265–278.

GREEN, R. K., S. MALPEZZI, AND S. K. MAYO (2005): "Metropolitan-Specific Estimates of the Price Elasticity of Supply of Housing, and Their Sources," *American Economic Review*, 95(2), 334–339.

GYOURKO, J. E., A. SAIZ, AND A. SUMMERS (2008): "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index," *Urban Studies*, 45(3), 1–35.

GYOURKO, J. E., AND T. SINAI (2003): "The Spatial Distribution of Housing-Related Ordinary Income Tax Benefits," *Real Estate Economics*, 31(4), 527–575.

HALL, P. G. (1982): *Great Planning Disasters*. University of California Press, Berkeley, CA, 2nd edn.

HERZOG, HENRY W, J., AND A. M. SCHLOTTMANN (1993): "Valuing Amenities and Disamenities of Urban Scale: Can Bigger be Better?," *Journal of Regional Science*, 2(2), 145–165.

KNOLL, M. S., AND T. D. GRIFFITH (2003): "Taxing Sunny Days: Adjusting Taxes for Regional Living Costs and Amenities," *Harvard Law Review*, 116(4), 987–1025.

KYDLAND, F. E., AND E. C. PRESCOTT (1977): "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3), 473–492.

MEADE, J. E. (1970): *The Theory of Indicative Planning: Lectures given in the University of Manchester*. Manchester University Press.

MILLER, J. B. (1980): "A Method for Determining the Appropriateness of National Planning in a Market Economy," *Quarterly Journal of Economics*, 95(2), 261–276.

NORDHAUS, W. D., AND J. TOBIN (1972): *The Measurement of Economic and Social Performance, Studies in Income and Wealth* vol. 38, chap. Is Growth Obsolete?, pp. 509–564. NBER, Cambridge, MA.

RAPPAPORT, J. (2008): "Consumption Amenities and City Population Density," *Regional Science and Urban Economics*, 38(3), 533–552.

RICHARDSON, H. W. (1972): "Optimality in City Size, Systems of Cities and Urban Policy: A Sceptic's View," *Urban Studies*, 9(1), 29–48.

——— (1973): "Theory of the Distribution of City Sizes: Review and Prospects," *Regional Studies*, 7(3), 239–251.

STOVER, M. E., AND C. L. LEVEN (1992): "Methodological Issues in the Determination of the Quality of Life in Urban Areas," *Urban Studies*, 29, 737–754.

WALSH, R. (2007): "Endogenous Open Space Amenities in a Locational Equilibrium," *Journal of Urban Economics*, 61(3), 319–344.

WILLIAMSON, J. G. (1981): "Urban Disamenities, Dark Satanic Mills, and the British Standard of Living Debate," *Journal of Economic History*, 41(1), 75–83.

Chapter 9

Jackknife instrumental variable estimation: A review

9.1 Abstract

Two-stage-least-squares (2SLS) estimates are biased towards the probability limit of OLS estimates. This bias grows with the degree of overidentification and can generate highly misleading results. Angrist, Imbens, and Krueger (1999) propose an alternative to 2SLS for models with more instruments than endogenous regressors, the so-called jackknife instrumental variables estimator (JIVE). This class of estimators is first-order equivalent to 2SLS but with superior finite-sample properties in term of bias and convergence rate of confidence intervals

Keywords: Instrumental variables, jackknife, bootstrap, Monte Carlo simulation
JEL Codes: C1, C2, C3

9.2 Introduction

Nagar (1959) used an approximation argument to show that two-stage least squares (2SLS) estimates are biased toward the probability limit of ordinary least squares (OLS) estimates in finite samples with normal disturbances. This bias grows with the degree of overidentification and can generate highly misleading results. The dispersion of the density decreases with the number of instruments from which one might (erroneously) conclude that there is a trade-off between bias and efficiency. However, contrary to conventional wisdom, there is no systematic relationship between the 2SLS bias and the number of instruments.

Angrist, Imbens, and Krueger (1999, AIK henceforth) propose an alternative to 2SLS for models with more instruments than endogenous regressors, the so-called jackknife instrumental variables estimator (JIVE). This class of estimators is first-order equivalent to 2SLS but with superior finite-sample properties in terms of bias and convergence rate of confidence intervals. In this context, Donald and Newey (2000) show that the continuous updating estimator can be interpreted as a jackknife estimator. This interpretation gives some insight as to why there is less bias associated with this estimator.

The remainder of the paper is structured as follows. Section 9.3 provides an overview of the literature on resampling techniques and bias correction for instrumental variables. Section 9.4 looks at the mechanics of the JIVE, whereas section 9.5 describes the data and experimental set-up in AIK. Section 9.6 replicates the Monte Carlo experiment in AIK and proposes a possible extension. A brief conclusion and directions for future research are discussed in section 9.7.

9.3 Literature review

Exact-finite sample results are not available for most econometric estimators and their related tests. As inference that relies on asymptotic theory is not practical in the context of most empirical work, resampling techniques in a Monte Carlo setting provide an attractive alternative approximation.

9.3.1 Bootstrap and jackknife for bias correction

The econometrics literature focuses on the use of resampling techniques in hypothesis testing, confidence intervals, estimation of standard error, and bias reduction. Bootstrap and its precursor the jackknife are popular resampling methods used for bias correction. Miller (1974) provides a comprehensive overview of the role of the jackknife technique in bias reduction.

The bootstrap as a resampling technique is due to Efron (1979) and Efron and Gong (1983) provide a comparison of the bootstrap, jackknife and other techniques that are used for the nonparametric estimation of "statistical error". Indeed, from an empirical perspective, one of the most appealing feature of these techniques is "[...] the substitution of raw computing power for theoretical analysis" (Efron and Gong, 1983, p.36). A detailed discussion of bootstrap inference in econometrics can be found in MacKinnon (2002).

9.3.2 Properties of IV estimators

The simple IV estimator

The interpretation of 2SLS as an IV estimator yields

$$\widehat{\boldsymbol{\beta}}_{2\text{SLS}} = \boldsymbol{\beta} + (\widehat{\mathbf{X}}'\mathbf{X})^{-1}\widehat{\mathbf{X}}'\boldsymbol{u}.$$

By assumption $n^{-1}\mathbf{Z}'\boldsymbol{u} = \mathbf{0}$ thus $\text{plim}(n^{-1}\widehat{\mathbf{X}}'\boldsymbol{u}) = \mathbf{0}$ and so $\widehat{\boldsymbol{\beta}}_{2\text{SLS}}$ is consistent. However, the reason for IV estimation in the first instance is endogeneity in some of the regressors, i.e. the correlation between $\mathbf{X}$ and $\boldsymbol{u}$, which means that $\widehat{\mathbf{X}} = \mathbf{P_Z X}$ is correlated with $\boldsymbol{u}$. Thus $\text{E}[\widehat{\mathbf{X}}'\boldsymbol{u}] \neq \mathbf{0}$ leads to the bias in the IV estimator which arises from using $\widehat{\mathbf{X}} = \mathbf{Z}\widehat{\boldsymbol{\pi}}$ rather than $\widehat{\mathbf{X}} = \mathbf{Z}\boldsymbol{\pi}$ as the instrument. An alternative is to use predictions $\widetilde{\mathbf{X}}$ as instruments, since these instruments have the property that $\text{E}[\widetilde{\mathbf{X}}'\boldsymbol{u}] = \mathbf{0}$ in addition to $\text{plim}(n^{-1}\widehat{\mathbf{X}}'\boldsymbol{u}) = \mathbf{0}$. This yields the estimator

$$\widetilde{\boldsymbol{\beta}}_{SIV} = (\widetilde{\mathbf{X}}'\mathbf{X})^{-1}\widetilde{\mathbf{X}}'\boldsymbol{u}.$$

However, this estimator will still be biased, since $\text{E}[\widetilde{\mathbf{X}}'\boldsymbol{u}] = \mathbf{0}$ does not imply $\text{E}[(\widetilde{\mathbf{X}}'\mathbf{X})^{-1}\widetilde{\mathbf{X}}'\boldsymbol{u}] = \mathbf{0}$.

Angrist and Krueger (1995) show that this bias may be reduced. They reevaluate their earlier IV estimates of the returns to schooling in light of the fact that 2SLS is biased in the same direction as OLS even in very large samples. They propose a split-sample IV estimator (SSIVE) which uses one-half of a sample to estimate parameters of the first-stage equation, and then using these parameter estimates to construct fitted values and second-stage parameter estimates in the other half-sample. The SSIVE is biased toward 0, but this bias can be corrected.

From SSIVE to JIVE

SSIVE incurs a considerable efficiency loss because only half the sample is used at the final stage. A more efficient variation on this theme would be to generate the instruments observation by observation. This is the main intuition behind the JIVE which was first proposed by Phillips and Hale (1977). The precise mechanics of JIVE are discussed in more detail in section 9.4.

Hahn and Hausman (2002a,b) derive a formula for the finite-sample bias of the 2SLS estimator which is proportional to the number of instrumental variables which implies a bias-variance trade-off. It may thus not be necessarily desirable to use all available instruments and a good estimator requires the appropriate selection of instruments.[1]

Also, Hahn, Hausmann, and Kürsteiner (2004) construct a bias-corrected version of 2SLS based on the Jackknife principle. Using higher-order expansions they show that the mean-squared error (MSE) of their Jackknife 2SLS estimator is approximately the same as the MSE of the Nagar-type estimator.

In one of the more extensive simulation studies Flores-Lagunes (2007) reports results comparing bias corrected versions of 2SLS, limited information maximum likelihood (LIML) and Fuller, a Jackknife version of 2SLS, and the REQML estimator, in settings with 100 and 500 observations, and 5 and 30 instruments for the single endogenous variable. He does not, however, include LIML with Bekker standard errors. He looks at median bias, median absolute error, inter decile range, coverage rates and concludes that "[...] that the random effects quasi-maximum likelihood estimator outperforms alternative estimators in terms of median point estimates and coverage rates." (Flores-Lagunes, 2007, p.692).

Weak instruments and finite sample bias

Standard normal asymptotic approximation to sampling distribution of IV, 2SLS, and LIML estimators relies on non-zero correlation between instruments and endogenous regressors. If correlation is close to zero, these approximations are not accurate even in fairly large samples. In the case with large degree of overidentification 2SLS has poor properties; considerable bias towards OLS, and substantial underestimation of standard errors. Comprehensive surveys on the literature of weak instruments can be found in Stock, Wright, and Yogo (2002) and Andrews and Stock (2005, 2007). Refer to Imbens (2007), and possibly Bekker (1994).

Furthermore, it is important to emphasise that if instruments are weak, even mild instrument endogeneity (i.e. invalid instrument) can lead to IV methods being *more inconsistent* than OLS. The asymptotic theory of weak instruments is developed comprehensively in Staiger and Stock (1997) and Mikusheva and Poi (2006), whereas Angrist and Pischke (2009) provide a series of illustrations of this problem in an applied setting.

There is a growing literature that addresses the practical problem that – with weak instruments – even in large samples, asymptotic theory provides a poor approximation to the distribution of IV estimators. While asymptotically consistent, IV estimators are biased towards the inconsistent OLS estimator in finite samples. This point is famously shown by Bound, Jaeger, and Baker (1995) using the AK91 data set that

[1] Asymptotic properties of bias correction for IV estimators are examined in Chao and Swanson (2007).

contains a large sample with over 300,000 observations.

Phillips (1980) derives the exact probability density function of instrumental variable estimators of the coefficients of the endogenous variables in a structural equation containing $n+1$ endogenous variables and N degrees of overidentification. Furthermore, since Phillips also notes that the dispersion of the density decreases with the number of instruments, one might conclude that there is a trade-off between bias and efficiency.[2] However, Buse (1992) demonstrates that – contrary to conventional wisdom – there is no systematic relationship between the 2SLS bias and the number of instruments.

Cameron and Trivedi (2005) highlight that finite-sample bias problems arise not only for the IV estimate, but also for IV standard errors and test statistics.

Recent developments in the "Case against JIVE"

Recently Blomquist and Dahlberg (1999, 2006) and Davidson and MacKinnon (2006, 2007) have questioned the appropriateness and validity of JIVE in empirical work. This is particularly due to the poor finite-sample properties of JIVE when instruments are very weak. I am addressing this point in my extension of AIK in section 9.6.3.

9.4 The mechanics of JIVE

9.4.1 The jackknife and bias correction

As an alternative to the bootstrap, the jackknife resampling method can be used for bias correction. Following the notation of Miller (1974), the jackknife is based on the idea of splitting the sample of size N into g groups of size h each. Then, for a sample of independent and identically distributed random variables, let $\hat{\theta}$ be an estimator of the parameter θ based on the sample size n. Let $\hat{\theta}_{-i}$ be the corresponding estimator based on the sample size $(g-1)h$, where the ith group of size h has been deleted. Then

$$\tilde{\theta}_i = g\hat{\theta} - (g-1)\hat{\theta}_{-i}, \quad \forall \ i = 1\ldots, g.$$

[2] This notion is based on the assumption that all the other parameters of the density are held fixed, including a noncentrality parameter. Buse (1992) highlights that the noncentrality parameter depends on the number of instruments.

The estimator

$$\tilde{\theta} = \frac{1}{g}\sum_{i=1}^{g}\tilde{\theta}_i = g\hat{\theta} - (g-1)\sum_{i=1}^{g}\hat{\theta}_i$$

has the property that it eliminates the order $1/N$ terms from the bias of the from $\mathrm{E}(\hat{\theta}) = \theta + a_n/n + O(1/n^2)$.

The jackknife estimate of the bias of $\hat{\theta}$ is $(n-1)(\bar{\hat{\theta}} - \hat{\theta})$, where $\bar{\hat{\theta}} = n^{-1}\sum_{i=1}^{n}\hat{\theta}_{-i}$ is the average of the n jackknife replications $\hat{\theta}_{-i}$. This leads to the bias-corrected jackknife estimate of θ, i.e.

$$\hat{\theta}_{\text{Jack}} = \hat{\theta} - (\bar{\hat{\theta}} - \hat{\theta}) = n\hat{\theta} - (n-1)\bar{\hat{\theta}}.$$

This reduces the bias from $O(n^{-1})$ to $O(n^{-2})$, which is the same order of bias reduction as for the bootstrap. However, because it uses less data, this estimator can have an increased variance compared with $\hat{\theta}$.

9.4.2 The bias of 2SLS

Basic linear model considered by AIK contains two equations. The first equation describes the economic relation of interest between an endogenous variable y_i and a vector of potentially endogenous regressors $\boldsymbol{x}_i$. The second equation captures the relation between the endogenous regressors and a vector of instruments $\mathbf{z}_i$.

$$\begin{aligned} y_i &= \boldsymbol{x}_i\boldsymbol{\beta} + u_i \\ \boldsymbol{x}_i &= \mathbf{z}_i\boldsymbol{\pi} + \boldsymbol{v}_i \end{aligned}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{v}_i$ are $1\times l$, $\mathbf{z}_i$ is $1\times k$, $\boldsymbol{\beta}$ is $1\times l$, and $\boldsymbol{\pi}$ is $k\times l$. By assumption, some of the regressors in $\boldsymbol{x}_i$ are correlated with $\boldsymbol{u}_i$, but all the elements of $\boldsymbol{z}_i$ are uncorrelated with $\boldsymbol{v}_i$. Let $\boldsymbol{\sigma}_{uv}$ denote the $1\times l$ vector of covariances between u_i and $\boldsymbol{v}_i$. Given this set-up, the number of overidentifying restrictions is therefore $k-l$. In matrix notation this model can then be written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{u} \tag{9.1}$$

$$\mathbf{X} = \mathbf{Z}\boldsymbol{\pi} + \boldsymbol{v} \tag{9.2}$$

A central assumption here is that all observations $\boldsymbol{x}_i, y_i$ and $\mathbf{z}_i$ are independent and identically distributed. In the following, I am using the notation of Poi (2006) rather than that of AIK. As highlighted above, the bias in the IV estimator which arises from using $\widehat{\boldsymbol{x}}_i = \mathbf{z}_i\widehat{\boldsymbol{\pi}}$ rather than $\widehat{\boldsymbol{x}}_i = \mathbf{z}_i\boldsymbol{\pi}$ as the instrument. If $\boldsymbol{\pi}$ were known, then the optimal IV estimator of $\boldsymbol{\beta}$ would be

$$\widehat{\boldsymbol{\beta}}_{\text{opt}} = (\boldsymbol{\pi}'\mathbf{Z}'\mathbf{X})^{-1}\boldsymbol{\pi}'\mathbf{Z}'\boldsymbol{y}.$$

$\mathbf{Z}\boldsymbol{\pi}$ is referred to as the *optimal instrument matrix*, since $\mathrm{E}(\widehat{\boldsymbol{\beta}}_{\text{opt}}) = \boldsymbol{\beta}$. However, $\boldsymbol{\pi}$ is not known and the 2SLS estimator uses the OLS estimate of $\boldsymbol{\pi}$, such that

$$\begin{aligned} \widehat{\beta}_{2\text{SLS}} &= (\widehat{\boldsymbol{\pi}}\mathbf{Z}'\mathbf{X})^{-1}\widehat{\boldsymbol{\pi}}'\mathbf{Z}\boldsymbol{y} \\ &= (\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\boldsymbol{y} \end{aligned} \tag{9.3}$$

where $\widehat{\boldsymbol{\pi}} = (\mathbf{Z}'\mathbf{Z}^{-1}\mathbf{Z}'\mathbf{X})$, i.e. the parameter estimate from the first stage of regressing the endogenous regressors $\mathbf{X}$ on the instruments $\mathbf{Z}$.

The limiting distributions of both $n^{-1/2}(\widehat{\boldsymbol{\beta}}_{\text{opt}} - \boldsymbol{\beta})$ and $n^{-1/2}(\widehat{\boldsymbol{\beta}}_{2\text{SLS}} - \boldsymbol{\beta})$ are normal with mean zero and variance $(\boldsymbol{\pi}\,\mathrm{plim}(n^{-1}\mathbf{Z}'\mathbf{Z})'\boldsymbol{\pi})^{-1}\boldsymbol{\sigma}_u^2$. This is central to the motivation of the AIK approach which is hinged on the observation that $\widehat{\boldsymbol{\beta}}_{\text{opt}}$ has much better small sample properties than $\widehat{\boldsymbol{\beta}}_{2\text{SLS}}$, even though the two estimators have the same asymptotic normal distribution. It is well known that increasing the number of instruments increases the bias of $\widehat{\boldsymbol{\beta}}_{2\text{SLS}}$.

This can be directly seen from the fact that the first-stage fitted values $\mathbf{Z}\widehat{\boldsymbol{\pi}}$ can be written as $\mathbf{P_Z X} = \mathbf{Z}\boldsymbol{\pi} + \mathbf{P_Z}\boldsymbol{v}$, where $\mathbf{P_Z}$ is the projection matrix $\mathbf{Z}'(\mathbf{Z}'\mathbf{Z})^{-\mathbf{1}}\mathbf{Z}$. The second component of these fitted values $\mathbf{P_Z}\boldsymbol{v}$ is correlated with $\boldsymbol{v}$ and thus also with $\boldsymbol{u}$. More formally, this can be shown as follows

$$\begin{aligned} \mathrm{E}(\widehat{\boldsymbol{\pi}}'\mathbf{z}_i'\boldsymbol{u}_i) &= \mathrm{E}\{\mathrm{E}(\widehat{\boldsymbol{\pi}}'\mathbf{z}_i' u_i)|\mathbf{Z}\} \\ &= \mathrm{E}\{\mathrm{E}[(\boldsymbol{\pi}'\mathbf{z}_i'\boldsymbol{u}_i + \boldsymbol{v}'\mathbf{z}_i(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}_i' u_i')|\mathbf{Z}\} \\ &= \mathrm{E}\{\mathbf{z}_i(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}_i' \cdot \mathrm{E}(u_i\boldsymbol{v}_i)\} \\ &= \frac{k}{n}\boldsymbol{\sigma}_{uv} \end{aligned} \tag{9.4}$$

AIK point out that even though this correlation disappears in large samples as $\frac{k}{n} \to 0$, it increases with the number of instruments, ceteris paribus. Equation (9.4) shows that the bias for the 2SLS estimator arises from the correlation between the OLS estimate of the optimal instrument matrix $\mathbf{z}_i\widehat{\boldsymbol{\pi}}$ and the residual $\boldsymbol{u}$.

9.4.3 Jackknife instrumental variable estimation

Both AIK and Blomquist and Dahlberg (1999) propose the JIVE as an alternative estimator of $\mathbf{z}_i\boldsymbol{\pi}$, since it is shown not to suffer from such correlation. Building on the work of Phillips and Hale (1977), AIK propose to use all observations except observation i to estimate the coefficient $\boldsymbol{\pi}$. This estimate is then used with $\mathbf{z}_i$ to compute the fitted value of the instrument for observation i. This process is repeated for each observation, each time obtaining a "leave-one-out" value of $\widehat{\boldsymbol{\pi}}$, which can be expressed as

$$\widetilde{\boldsymbol{\pi}}_{-i} = (\mathbf{Z}'_{-i}\mathbf{Z}_{-i})^{-1}\mathbf{Z}'_{-i}\mathbf{X}_{-i} \tag{9.5}$$

where $\mathbf{Z}_{-i}$ and $\mathbf{X}_{-i}$ are the $(n-1)\times k$ and $(n-1)\times l$ matrices equal to $\mathbf{Z}_{-i}$ and $\mathbf{X}_{-i}$ respectively with the ith row removed. The ith row estimate of the optimal instrument is thus $\mathbf{z}_i\widehat{\boldsymbol{\pi}}_{-i}$. Key to the argument of why the finite-sample properties of JIVE are argued to be superior to those of 2SLS is the following property. Because u_i is independent of $\boldsymbol{x}_j$ if $i \neq j$, it is then the case that

$$\begin{aligned}
\mathrm{E}(\hat{\boldsymbol{\pi}}'_{-i}\mathbf{z}'_i u_i) &= \mathrm{E}\{(\mathbf{X}'_{-i}\mathbf{Z}_{-i}(\mathbf{Z}'_{-i}\mathbf{Z}_{-i})^{-1}\mathbf{z}'_i u_i)|\mathbf{Z}\} \\
&= \mathrm{E}\{\mathbf{z}_i(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}'_i \cdot \mathrm{E}(\mathbf{X}'_i u_i|\mathbf{Z})\} \\
&= \mathbf{0}
\end{aligned} \tag{9.6}$$

Thus, we obtain an unbiased estimator for $\boldsymbol{\beta}$. This version of the JIVE is defined by

$$\widehat{\boldsymbol{\beta}}_{\text{JIVE1}} = (\hat{\mathbf{X}}'\mathbf{X})^{-1}\hat{\mathbf{X}}'\boldsymbol{y},$$

where the ith row of $\hat{\mathbf{X}}$ is given by $\mathbf{z}_i\widehat{\boldsymbol{\pi}}$. In contrast to equation (2.2), JIVE1 does not suffer from the finite-sample bias of 2SLS due to equation (A-1). In addition, $\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}$ is also a consistent estimator of $\boldsymbol{\beta}$ because

$\widehat{\boldsymbol{\pi}}_{-i}$ is a consistent estimator of $\boldsymbol{\pi}$.

AIK highlight that for empirical purposes it is not necessary to calculate the n regression coefficients $\widehat{\boldsymbol{\pi}}_{-i}$. Instead, an empirical implementation of JIVE only requires the construction of the instrument which is calculated as

$$\begin{aligned}
\mathbf{z}_i \widehat{\boldsymbol{\pi}}_{-i} &= \mathbf{z}_i \frac{(\mathbf{Z}'\mathbf{Z})^{-1}}{1-\mathbf{z}_i(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}_i'}(\mathbf{Z}'\mathbf{X}-\mathbf{z}_\mathbf{i}'\boldsymbol{x}_\mathbf{i}) \\
&= \frac{\mathbf{z}_i\widehat{\boldsymbol{\pi}}_i - h_i \boldsymbol{x}_i}{1-h_i} \qquad (9.7)
\end{aligned}$$

where the hat matrix $h_i = \mathbf{z}_i(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{z}_i'$.[3] Thus the calculation of $\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}$ can be achieved in two steps. First, the regular fitted first-stage values and the leverage h_i are computed, then – as a second step – the $\mathbf{z}_i\widehat{\boldsymbol{\pi}}_{-i}$ are used as instrument.

The variance of $\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}$ is given by

$$\text{Var}(\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}) = (\widehat{\mathbf{X}}'\mathbf{X})^{-1}\widehat{\mathbf{X}}'\boldsymbol{u}\boldsymbol{u}'\widehat{\mathbf{X}}(\mathbf{X}'\widehat{\mathbf{X}})^{-1}.$$

Under homoskedastic errors $\boldsymbol{u}$, the estimator of this covariance matrix is simply

$$\widehat{\text{Var}}(\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}) = \widehat{\boldsymbol{\sigma}^2}(\widehat{\mathbf{X}}'\mathbf{X})^{-1}\widehat{\mathbf{X}}'\widehat{\mathbf{X}}(\mathbf{X}'\widehat{\mathbf{X}})^{-1},$$

where

$$\widehat{\boldsymbol{\sigma}^2} = \frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{y}_i - \boldsymbol{x}_i\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}\right)^2.$$

However, Chao, Swanson, Hausmann, Newey, and Woutersen (2007) derive the asymptotic distribution of JIVE if $\boldsymbol{u}$ displays heteroskedasticity and a heteroskedasticity-robust estimator is defined as

$$\widehat{\text{Var}}(\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}) = (\widehat{\mathbf{X}}'\mathbf{X})^{-1}\sum_{i=1}^{n}\hat{u}_i^2\widehat{\boldsymbol{x}}_i\widehat{\boldsymbol{x}}_i'(\mathbf{X}'\widehat{\mathbf{X}})^{-1},$$

where

$$\hat{u}_i^2 = \frac{1}{n}\left(y_i - \boldsymbol{x}_i\widehat{\boldsymbol{\beta}}_{\text{JIVE1}}\right)^2.$$

[3] See Cook (1979) for a discussion of leverage in the context of influential observations.

9.4.4 Variations on the theme

JIVE1 removes the dependence of the constructed instrument by using a “leave-one-out” bootstrapped fitted value instead of the standard first stage regression. AIK also propose a variation on the theme, JIVE2, whereby only the $\mathbf{Z}'\mathbf{X}$ component of $\widehat{\boldsymbol{\pi}} = (\mathbf{Z}'\mathbf{Z})^{-1}(\mathbf{Z}'\mathbf{X})$ is adjusted. Thus JIVE2 is also based on the idea of eliminating the correlation between the ith observation of the constructed instrument and the ith observation of the endogenous regressor, but now $\widetilde{\boldsymbol{\pi}}_{-i}$ is redefined as

$$\widetilde{\boldsymbol{\pi}}_{-i} \quad = \quad (\mathbf{Z}'_{-i}\mathbf{Z}_{-i})^{-1}\mathbf{Z}'_{-i}\mathbf{X}_{-i} \tag{9.8}$$

Therefore,

$$\widehat{\boldsymbol{\beta}}_{\text{JIVE2}} = (\widehat{\mathbf{X}}'\mathbf{X})^{-1}\widehat{\mathbf{X}}'\boldsymbol{y}.$$

In addition, Blomquist and Dahlberg (1999) also propose the estimator

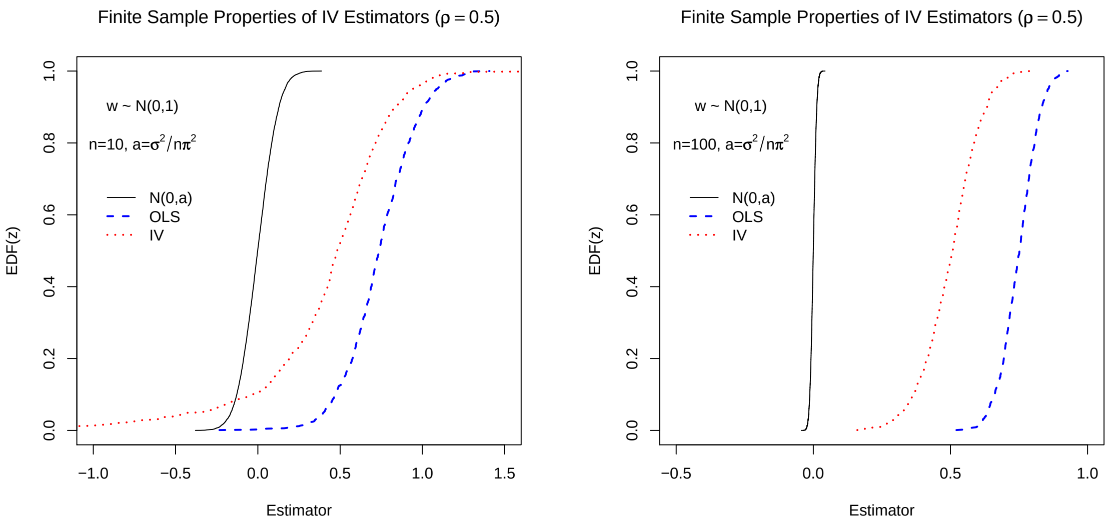
$$\widehat{\boldsymbol{\beta}}_{\text{JIVE3}} = (\widehat{\mathbf{X}}'\widehat{\mathbf{X}})^{-1}\widehat{\mathbf{X}}'\boldsymbol{y},$$

which uses $\widehat{\boldsymbol{\pi}}_{-i}$ as in equation (9.5). However, while JIVE3 has the slight advantage of only requiring OLS once $\widehat{\mathbf{X}}$ has been computed, this estimator is not unbiased.

9.5 Finite sample properties of IV estimators

In contrast to the asymptotic properties, the finite sample distributions of IV estimators are – with the exception of special cases – not known in practice, which renders finite-sample inference subject to serious error. In the just identified case, the IV estimator does not have any moments as the tails of its distribution are so thick that its expectation does not exists. Nonetheless, the IV is estimator is consistent and the number of available moments of its distribution rises with the degree of overidentification.

Figure 9.1: IV finite sample properties: Impact of sample size on CDF



(a) Small sample ($n = 10$)

(b) Larger sample ($n = 100$)

9.5.1 Simulation of IV finite-sample properties

Put differently, in the overidentified case, the first $l - k$ moments of $\hat{\boldsymbol{\beta}}_{\text{IV}}$ exist, which is consistent with the result for the just identified model, where $l - k = 0$ and the IV estimator does not have any moments.[4] These properties are visualised in the following section using the Monte Carlo set-up of AIK.

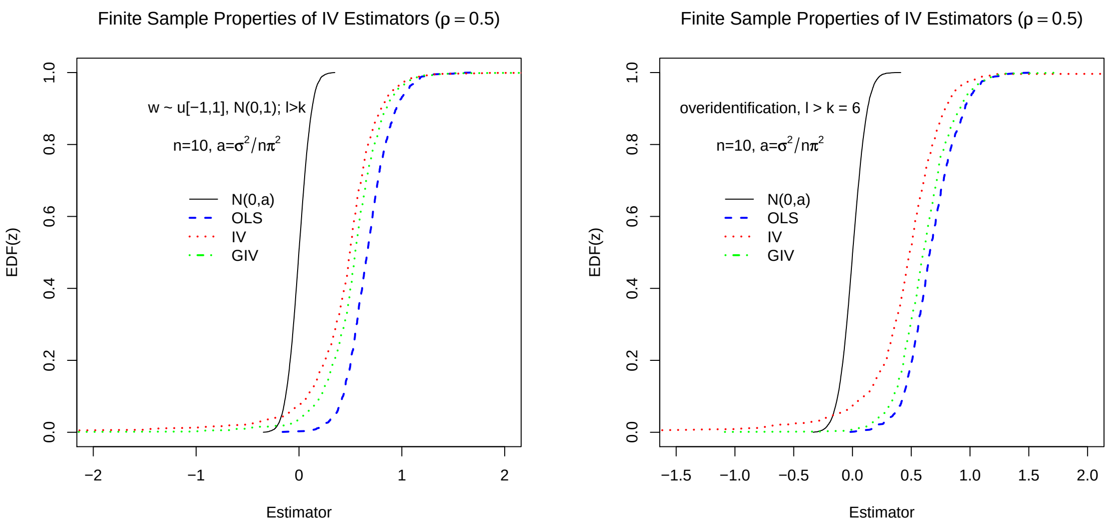
Figure 9.1 highlights the impact of increasing the sample from $n = 10$ to $n = 100$; as the sample size increases, the tails of the empirical distribution flatten relatively quickly and thus convergence to the asymptotic distribution (i.e. the normal distribution) happens already in relatively small samples. Increasing the sample size further to $n = 1{,}000$ yields an even "steeper" EDF, although the graph is not included here as it was too big (~20MB).

It is important to emphasise that different instrument DGPs do not have an impact on the empirical distribution of $\hat{\beta}_{\text{IV}}$. The main reason is due to the fact that the IV estimator depends on $\mathbf{Z}$ only through the orthogonal projection matrix $\mathbf{P_Z}$. Thus all that matters is the space $\mathscr{S}(\mathbf{Z})$ spanned by the instruments,

[4] Davidson and MacKinnon (2004) highlight that this does not contradict the result that $\hat{\boldsymbol{\beta}}_{\text{IV}}$ is asymptotically normal. Specifically, even if a set of random variables converges to a limiting random variable, the moments of these random variables do not necessarily converge to those of the limiting variable

not the actual DGP of the instruments.

Figure 9.2: IV finite sample properties: Impact of overidentification on CDF



(a) Weak overidentification

(b) Strong overidentification

Figure 9.2 highlights the impact of overidentification on the empirical distribution of $\hat{\beta}_{\mathrm{IV}}$. As the level of overidentification increases from panel 9.2(a) to panel (a) – i.e. identification rises from $l - k = 2$ to $l - k = 6$ – the bias increases while the tails of the EDF become thinner. Thus the theoretical prediction that the overidentified GIV estimator is more biased (but less leptokurtic) than the just-identified Simple IV estimator is verified.

9.6 Replication

This section replicates and extends the Monte Carlo simulation of AIK. As the first step, the finite sample properties of JIVE1 and JIVE2 are investigated for five different models and evaluated against the performance of OLS and 2SLS.

9.6.1 Different DGPs

AIK begin with a model where there is a single overidentifying restriction. The second model is similar, with the modification that there are a large number of instruments relative to the number of regressors. In the third model, the first stage is non-linear and heteroscedastic. The fourth model sets the coefficients to zero for all instruments and model 5 contains misspecification where one of the instruments has incorrectly been left out of the main regression.

Model 1

This is the most basic IV specification with a single overidentifying restiction such that the JIVE has at least one moment.[5]

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + u_i \\ x_{i1} &= \pi_0 + \sum_{j=1}^{2} \pi_j z_{ij} + v_i \end{aligned}$$

with $\beta_0 = 0$, $\beta_1 = 1$, $\pi_0 = 0$, $\pi_1 = 0.3$, and $\pi_2 = 0$. Here, $k = 3$ and $l = 2$, and

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \Big| \mathbf{Z} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.25 & 0.20 \\ 0.20 & 0.25 \end{pmatrix} \right).$$

All z_{ij} are independent, normally distributed random variables with mean zero and unit variance. The contemporaneous correlation of $\boldsymbol{u}$ and $\boldsymbol{v}$ implies $\mathrm{E}(u_t v_t)$ for some correlation coefficient ρ such that $-1 < \rho < 1$. Although not explicitly stated in AIK, the covariance matrix in model 1

$$\begin{pmatrix} 0.25 & 0.20 \\ 0.20 & 0.25 \end{pmatrix} = \begin{pmatrix} 0.25 & \rho \cdot \sqrt{0.5}\sqrt{0.5} \\ \rho \cdot \sqrt{0.5}\sqrt{0.5} & 0.25 \end{pmatrix}$$

implies a correlation coefficient of $\rho = 0.8$, i.e. model 1 assumes strong instruments. This assumption is also maintained for the remaining models. As an extension to AIK, I will investigate the impact to weak instruments on the properties of JIVE, rather than assuming the inclusion of invalid instruments as is the case in models 2,3 and 5.

[5]As highlighted above, the IV estimator does not have any moments in the just-identified case.

Model 2

Model 2 adds 18 worthless instruments to the design in Model 1, à la Bound, Jaeger, and Baker (1995):

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + u_i \\ x_{i1} &= \pi_0 + \sum_{j=1}^{20} \pi_j z_{ij} + v_i \end{aligned}$$

with $\beta_0 = 0$, $\beta_1 = 1$, $\pi_0 = 0$, $\pi_1 = 0.3$, and $\pi_j = 0$ for $j = 2, 3, \ldots, 20$. Here, $k = 21$ and $l = 2$, and

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \Big| \mathbf{Z} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.25 & 0.20 \\ 0.20 & 0.25 \end{pmatrix} \right).$$

All z_{ij} are independent, normally distributed random variables with mean zero and unit variance.

Model 3

Model 3 has the same basic structure as above, except that the relationship between $\boldsymbol{x}_i$ and $\mathbf{z}_i$ is non-linear and heteroskedastic. As in model 2, there are 20 linear instruments and non-linearities are ignored in the first stage:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + u_i \\ x_{i1} &= \pi_0 + \sum_{j=1}^{20} \pi_j z_{ij} + 0.3 \sum_{j=2}^{20} z_{ij}^2 + v_{i0} + \sum_{j=2}^{20} z_{ij}^2 / 19 \end{aligned}$$

with $\beta_0 = 0$, $\beta_1 = 1$, $\pi_0 = 0$, $\pi_1 = 0.3$, and $\pi_j = 0$ for $j = 2, 3, \ldots, 20$. Here, $k = 21$ and $l = 2$, and

$$\begin{pmatrix} u_i \\ v_{i0} \end{pmatrix} \Big| \mathbf{Z} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1.0 & 0.8 \\ 0.8 & 1.0 \end{pmatrix} \right).$$

All z_{ij} are independent, normally distributed random variables with mean zero and unit variance.

Model 4

Same basic structure as model 2, but with all coefficients in the reduced from set to zero:

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + u_i \\ x_{i1} &= \pi_0 + \sum_{j=1}^{20} \pi_j z_{ij} + v_i \end{aligned}$$

with $\beta_0 = 0$, $\beta_1 = 1$, $\pi_0 = 0$, $\pi_1 = 0.3$, and $\pi_j = 0$ for $j = 2, 3, \ldots, 20$. Here again, $k = 21$ and $l = 2$, and

$$\begin{pmatrix} u_i \\ v_i \end{pmatrix} \Big| \mathbf{Z} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.25 & 0.20 \\ 0.20 & 0.25 \end{pmatrix} \right).$$

Model 5

Same basic structure as model 2, but one of the instruments that has a zero coefficient in the second regression has a non-zero coefficient in the first regression which is inappropriately left out.

$$\begin{aligned} y_i &= \beta_0 + \beta_1 x_{i1} + u_i \\ x_{i1} &= \pi_0 + \sum_{j=1}^{20} \pi_j z_{ij} + v_i \end{aligned}$$

with $\beta_0 = 0$, $\beta_1 = 1$, $\pi_0 = 0$, $\pi_1 = 0.3$, and $\pi_j = 0$ for $j = 2, 3, \ldots, 20$. Here again, $k = 21$ and $l = 2$. But now, $u_i = 0.2 z_{i2} + \epsilon_i$ and

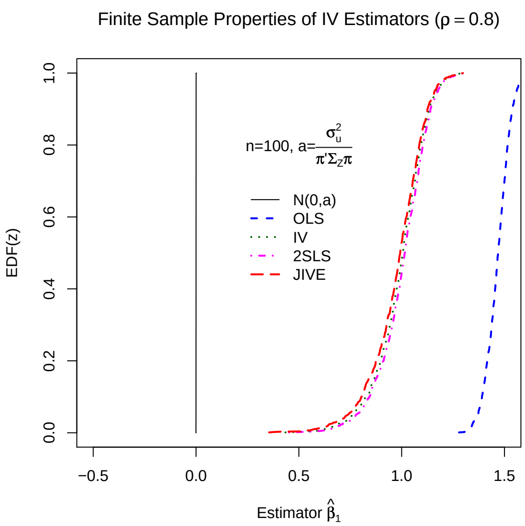
$$\begin{pmatrix} \epsilon_i \\ v_i \end{pmatrix} \Big| \mathbf{Z} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.25 & 0.20 \\ 0.20 & 0.25 \end{pmatrix} \right).$$

This means that because y_i directly depends on z_{i2}, the instrument z_{i2} is not valid. However, because x_i does not depend on z_{i2}, the estimated instrument does also not depend on z_{i2} and therefore both 2SLS and JIVE are not affected by this misspecification.

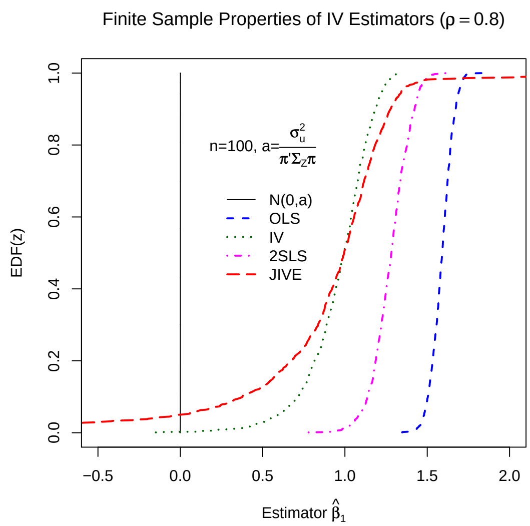
9.6.2 Monte Carlo simulation

This section will replicate AIK, using the models for 5 different DGPs. It might also be interesting to explore how the results react to different MC model specifications. The results from models 1,2,4 and 5 are displayed in table 9.1. The most striking result from these simulations is fact that – given a certain type of instrument strengh – JIVE seems to perform relatively well across different model specifications. Here I follow AIK and assume that instruments are relatively weak, taking the empirical correlations from Angrist and Krueger (1991, AK91). This is later then relaxed and instruments are assumed to be even weaker. I am thus able to replicate the main result of AIK. However, exploratory computations in the next section investigate the stability of these results. The simulation results for models 1,2,4 and 5 are visualised in figure 9.3.

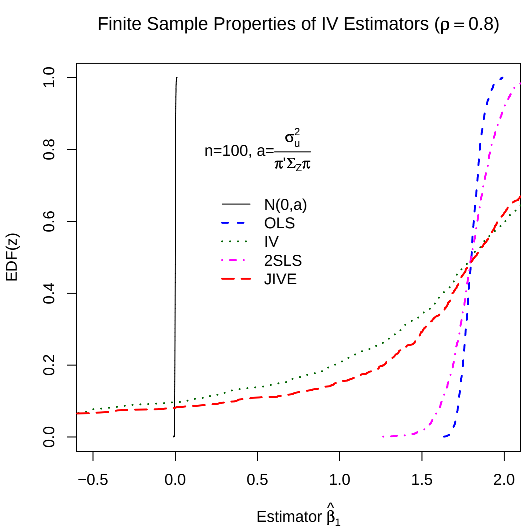
Figure 9.3: JIVE AIK simulations: Impact of model specification on CDF



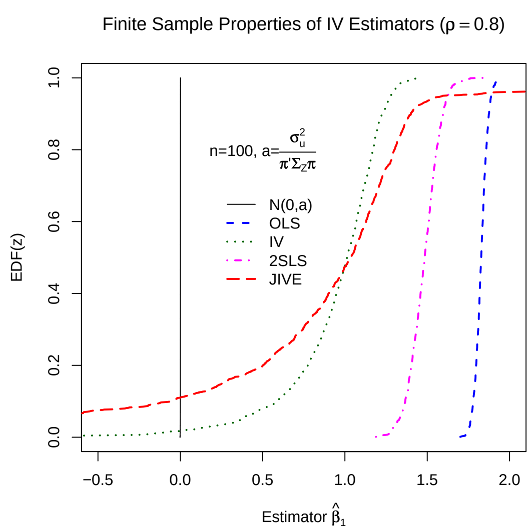
(a) Model 1



(b) Model 2



(c) Model 4



(d) Model 5

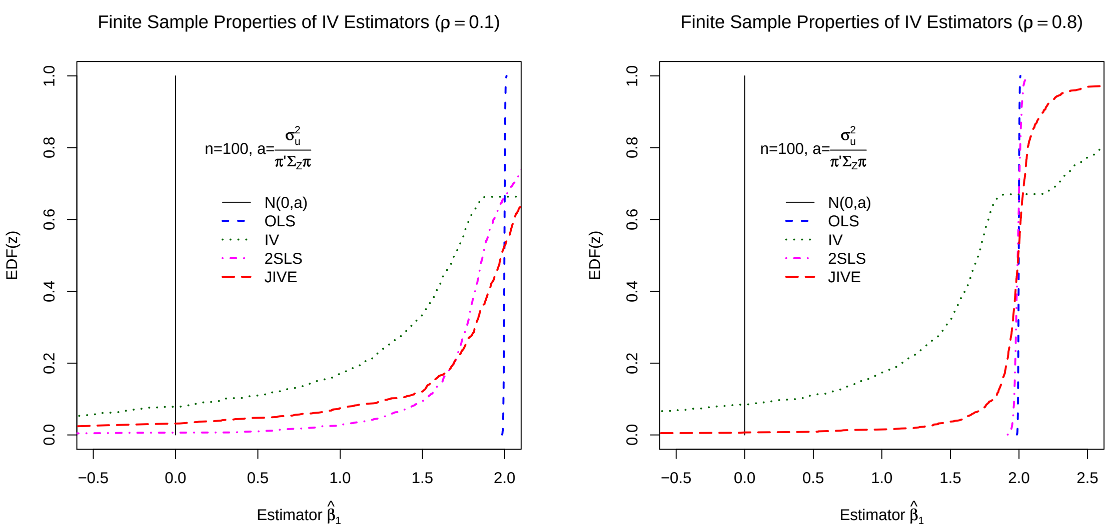
Table 9.1: Replication results from AIK with AK91 instruments ($\pi_1 = 0.3$)

Estimator	Quantiles around β_1					Median
	0.10	0.25	0.50	0.75	0.90	abs. error
	(1)	(2)	(3)	(4)	(5)	(6)
Model 1: $N = 100, k = 2, l = 3$; 5000 replications						
OLS	1.3891	1.4237	1.4649	1.5032	1.5373	0.0400
IV	0.8269	0.9179	1.0038	1.0742	1.1339	0.0774
2SLS	0.8435	0.9298	1.013	1.0839	1.1428	0.0768
JIVE1	0.8231	0.9149	1.0036	1.0767	1.1370	0.0798
JIVE2	0.8228	0.9153	1.0036	1.0771	1.1372	0.0802
Model 2: $N = 100, k = 2, l = 21$; 5000 replications						
OLS	1.5038	1.5444	1.5904	1.634	1.6733	0.0447
IV	0.7151	0.8625	0.9951	1.0975	1.1863	0.1165
2SLS	1.1370	1.2044	1.2825	1.3518	1.4232	0.0750
JIVE1	0.3782	0.7686	0.9953	1.1508	1.2766	0.1762
JIVE2	0.4694	0.7786	0.9843	1.1337	1.2487	0.1716
Model 3: $N = 100, k = 2, l = 21$; 5000 replications						
OLS	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000
IV	0.9994	0.9998	1.0001	1.0002	1.0007	0.0002
2SLS	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000
JIVE1	0.9999	1.0000	1.0000	1.0000	1.0002	0.0000
JIVE2	0.9999	1.0000	1.0000	1.0000	1.0000	0.0000
Model 4: $N = 100, k = 2, l = 21$; 5000 replications						
OLS	1.7200	1.7578	1.7991	1.8394	1.8777	0.0410
IV	0.1408	1.2051	1.7667	2.3826	3.6481	0.5902
2SLS	1.6099	1.6958	1.7934	1.8896	1.9765	0.0972
JIVE1	0.1968	1.3543	1.7969	2.2242	3.1096	0.4375
JIVE2	0.1481	1.3657	1.7883	2.205	2.9833	0.4188
Model 5: $N = 100, k = 2, l = 21$; 5000 replications						
OLS	1.7793	1.8032	1.8280	1.8515	1.8751	0.0243
IV	0.5852	0.8301	1.0127	1.1468	1.2283	0.1497
2SLS	1.3665	1.418	1.4847	1.5422	1.596	0.0617
JIVE1	-0.0558	0.6333	1.0315	1.2524	1.4008	0.2770
JIVE2	-0.0117	0.6292	0.9822	1.2188	1.3569	0.2708

9.6.3 Extension

In addition to the AIK Monte Carlo study (which assumes the instrument correlations in AK91), I also perform a set of simulations that assumes that instruments are even weaker (i.e. $\pi_1 = 0.03$). The results for models 1 and 2 (the largely overidentified case) are presented in table 9.2. Interestingly, my results are very similar to the findings of Davidson and MacKinnon (2006) who find no evidence to suggest that JIVE should ever be used. It is always more dispersed than 2SLS, often very much so, and – perhaps most importantly – JIVE seems to perform particularly badly when the instruments are very weak.

Figure 9.4: IV finite sample properties: Impact of very weak instruments on CDF



(a) Model 1

(b) Model 2

The impact of weak instruments on JIVE is also illustrated in figure 9.4 which highlights that JIVE seems to perform particularly badly when the instruments are weak.

9.7 Conclusions and outlook

The search for appropriate instruments in finite samples lies at the heart of the recent interest in simulation-based studies that investigate the properties of a variety of IV estimators. The replication of AIK with an extension has shown that the properties of JIVE depend critically on the assumptions about the nature

Table 9.2: Results with very weak instruments ($\pi_1 = 0.03$)

Estimator	Quantiles around β_1					Median
	0.10	0.25	0.50	0.75	0.90	abs. error
	(1)	(2)	(3)	(4)	(5)	(6)
Model 1: $N = 100, k = 2, l = 3$; 5000 replications						
OLS	1.9926	1.9957	1.9984	2.0013	2.0038	0.0028
IV	0.5317	1.3108	1.6964	2.4596	3.2756	0.6066
2SLS	1.5159	1.7206	1.8704	2.1324	2.3807	0.1961
JIVE1	1.3684	1.7454	1.9966	2.2128	2.6247	0.2329
JIVE2	1.3678	1.7489	1.9979	2.2130	2.6164	0.2307
Model 2: $N = 100, k = 2, l = 21$; 5000 replications						
OLS	1.9928	1.9952	1.9983	2.0009	2.0036	0.0028
IV	0.3841	1.3273	1.7077	2.4044	3.5248	0.5633
2SLS	1.9644	1.9759	1.9907	2.0047	2.0186	0.0146
JIVE1	1.8243	1.9307	1.9944	2.0506	2.1788	0.0576
JIVE2	1.7951	1.9265	1.9963	2.0570	2.2014	0.0654

of the instruments. While JIVE seems relatively stable across different model specifications and is indeed superior to 2SLS in terms of bias, very weak instruments has a dramatic impact on the suitability of JIVE.

9.7.1 Further research

Further research could look at the following areas to see to what extent the results on JIVE with weak instruments changes.

How many bootstraps?

Davidson and MacKinnon (2000) argue that there is a loss of power due to bootstrapping with finite B, (integer rule $\alpha(B+1) \equiv$ integer). This could be compared to the 3-step numerical method proposed by Andrews and Buchinsky (2000) in this section. This method determines the number of B needed to ensure a given level of accuracy.

How many instruments?

Properties of instrumental variable estimators are sensitive to the choice of valid instruments, even in large cross-section applications. Donald and Newey (1999, 2001, DN01) and explore applicability to AIK, i.e. comparison between performance of bias-adjusted 2SLS (B2SLS) and JIVE, both with the DN01 instrument selection method. Okui (2005) shows that using an analytic approximation of the bias rather than the bootstrap bias estimate, the bootstrap-based procedure improves the performance of the 2SLS estimator, often better than the selection method of DN01.

With the current resurgence of interest in properties of IV estimators, the upshot of it all is that there seems an increasing consensus that JIVE should never be used – i.e. there are many reasons why it is bad, but there are other estimators that have the same good low bias properties but are better in many other respects. Recent work by Ackerberg and Devereux (2009) introduce two simple new variants of the jackknife instrumental variables estimator for overidentified linear models and show that they are superior to the existing JIVE estimator, significantly improving on its small-sample-bias properties. They also show that, in models with heteroskedasticity, this improved JIVE variant has superior properties compared to the bias-adjusted 2SLS (B2SLS) estimator introduced by DN01.

A comprehensive, more systematic comparison of the large and small-sample properties and exploring the similarity of the bias inherent in JIVE with the one that characterizes the finite-sample bias of non-parametric regression is left for future work.

9.A Acknowledgements

I am grateful to Chris Parmeter for valuable comments and suggestions on earlier drafts. I would also like to thank Stephen Donald for suggestions and a copy of the unpublished mimeograph "Choosing the Number of Instruments".

9.B References

ACKERBERG, D. A., AND P. J. DEVEREUX (2009): "Improved JIVE Estimators for Overidentified Linear Models with and without Heteroskedasticity," *Review of Economics and Statistics*, 91(2), 351–362.

ANDREWS, D. W. K., AND M. BUCHINSKY (2000): "A Three-Step Method for Choosing the Number of Bootstrap Repetitions," *Econometrica*, 68(1), 23–51.

ANDREWS, D. W. K., AND J. H. STOCK (2005): "Inference with Weak Instruments," Discussion Paper 1530, Cowles Foundation, New Haven, Connecticut.

——— (2007): "Testing with Many Weak Instruments," *Journal of Econometrics*, 138, 24–46.

ANGRIST, J. D., G. W. IMBENS, AND A. B. KRUEGER (1999): "Jackknife Instrumental Variables Estimation," *Journal of Applied Econometrics*, 14(1), 57–67.

ANGRIST, J. D., G. W. IMBENS, AND D. B. RUBIN (1996): "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91(434), 444–455.

ANGRIST, J. D., AND A. B. KRUEGER (1991): "Does Compulsory Schooling Attendance Affect Schooling and Earnings?," *Quarterly Journal of Economics*, 106(4), 979–1014.

——— (1995): "Split Sample Instrumental Variables Estimates of the Returns to Schooling," *Journal of Business & Economic Statistics*, 13(2), 225–235.

——— (2001): "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 15(4), 69–85.

ANGRIST, J. D., AND J.-S. PISCHKE (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, NJ.

BEKKER, P. A. (1994): "Alternative Approximations to the Distributions of Instrumental Variable Estimators," *Econometrica*, 62(3), 657–681.

BLOMQUIST, S., AND M. DAHLBERG (1999): "Small Sample Properties of LIML and Jackknife IV Estimators: Experiments with Weak Instruments," *Journal of Applied Econometrics*, 14(1), 69–88.

——— (2006): "The Case against JIVE: A Comment," *Journal of Applied Econometrics*, 21(3), 839?–841.

BOUND, J., D. A. JAEGER, AND R. M. BAKER (1995): "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable Is Weak," *Journal of the American Statistical Association*, 90(430), 443–450.

BUSE, A. (1992): "The Bias of Instrumental Variables Estimators," *Econometrica*, 60(1), 173–180.

CAMERON, A. C., AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press, New York, NY.

CHAO, J. C., AND N. R. SWANSON (2007): "Alternative Approximations of the Bias and MSE of the IV Estimator under Weak Identification with Application in Bias Correction," *Journal of Econometrics*, 137(2), 515?–555.

CHAO, J. C., N. R. SWANSON, J. A. HAUSMANN, W. K. NEWEY, AND T. WOUTERSEN (2007): "Asymptotic Distribution of JIVE in a Heteroskedastic IV Regression with Many Instruments," Mimeograph, University of Maryland, Rutgers University, MIT, and Johns Hopkins University.

COOK, R. D. (1979): "Influential Observations in Linear Regressions," *Journal of the American Statistical Association*, 74(1), 169–174.

DAVIDSON, R., AND J. G. MACKINNON (2000): "Bootstrap Tests: How Many Bootstraps?," *Econometric Reviews*, 19(1), 55–68.

——— (2004): *Econometric Theory and Methods*. Oxford University Press, 1st edn.

——— (2006): "The Case Against JIVE," *Journal of Applied Econometrics*, 21(3), 827?–833.

——— (2007): "Moments of IV and JIVE estimators," *Econometrics Journal*, 10(4), 541?–553.

DONALD, S. G., AND W. K. NEWEY (1999): "Choosing the Number of Instruments," Mimeograph, Boston University and MIT.

——— (2000): "A Jackknife Interpretation of the Continuous Updating Estimator," *Economics Letters*, 67(3), 239–243.

——— (2001): "Choosing the Number of Instruments," *Econometrica*, 69(5), 1161–1191.

EFRON, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7(1), 1–26.

EFRON, B., AND G. GONG (1983): "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation," *American Statistician*, 37(1), 36–48.

FLORES-LAGUNES, A. (2007): "Finite Sample Evidence of IV-Estimators under Weak Instruments," *Journal of Applied Econometrics*, 22(3), 677–?694.

HAHN, J., AND J. A. HAUSMAN (2002a): "A New Specification Test for the Validity of Instrumental Variables," *Econometrica*, 70(1), 163–189.

——— (2002b): "Notes on Bias in Estimators for Simultaneous Equation Models," *Economics Letters*, 75, 237–?241.

HAHN, J., J. A. HAUSMANN, AND G. KÜRSTEINER (2004): "Estimation with Weak Instruments: Accuracy of Higher-order Bias and MSE Approximations," *Econometrics Journal*, 7(1), 272?–306.

IMBENS, G. W. (2007): "Weak Instruments and Many Instruments," in *What?s New in Econometrics*, NBER Summer Institute, Cambridge, MA. NBER.

MACKINNON, J. G. (2002): "Bootstrap Inference in Econometrics," *Canadian Journal of Economics*, 35(4), 615–645.

MIKUSHEVA, A., AND B. P. POI (2006): "Tests and Confidence Sets with Correct Size in the Simultaneous Equations Model with Potentially Weak Instruments," *Stata Journal*, 6(3), 1–11.

MILLER, R. G. (1974): "The Jackknife – A Review," *Biometrika*, 61(1), 1–15.

NAGAR, A. L. (1959): "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations," *Econometrica*, 27(4), 575–595.

OKUI, R. (2005): "Bootstrap?-based Criteria for Choosing the Number of Instruments," in *MODSIM 2005 International Congress on Modelling and Simulation*, ed. by A. Zerger, and R. Argent. Modelling and Simulation Society of Australia and New Zealand.

PHILLIPS, G., AND C. HALE (1977): "The Bias of Instrumental Variable Estimators of Simultaneous Equation Systems," *International Economic Review*, 18(1), 219–228.

PHILLIPS, P. (1980): "The Exact Distribution of Instrumental Variable Estimators in an Equation Containing n + 1 Endogenous Variables," *Econometrica*, 48(4), 861–878.

POI, B. P. (2006): "Jackknife Instrumental Variables Estimation in Stata," *Stata Journal*, 6(3), 364?–376.

STAIGER, D., AND J. H. STOCK (1997): "Instrumental Variable Regressions with Weak Instruments," *Econometrica*, 65(3), 557–586.

STOCK, J. H., J. H. WRIGHT, AND M. YOGO (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of the American Statistical Association*, 20(4), 518–526.