Essays on Agricultural and Regional Development

Zhen Cheng

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
Agricultural and Applied Economics

Jeffrey Alwang, Chair
Suqin Ge
George Norton
Catherine Larochelle

June 24th, 2019
Blacksburg, VA

Keywords: Food demand in West Africa, staple foods, returns to agricultural research and extension, model uncertainty

Essays on Agricultural and Regional Development

Zhen Cheng

ABSTRACT

In a world of imbalance, food consumption exhibits great diversity among regions and countries. Although farmers in developed economies benefit from up-to-date agricultural technology and produce much more than they consume, households in the developing world are still combating food insecurity. This dissertation is composed of two manuscripts. One is about consumption in developing countries, while the other is related to promoting agricultural production in a developed economy.

Chapter 1 applies a three-stage demand system to nationally representative household survey data to identify food demand behavior with an emphasis on food staples in two West Africa countries – Niger and Nigeria. The third stage of the demand system offers demand elasticities of specific staple items. Instead of treating the population as a whole, the study distinguishes rural and urban households and households of different welfare status. Results confirm the complexity of the food and staples demand between rural/urban areas and among welfare quintiles. Therefore, researchers and policymakers should consider not only the average demand response but also its distribution among households. In addition to demand elasticities, the effects of household demographic characteristics on the structure of food consumption are also obtained.

Chapter 2 estimates the rates of return to Virginia's public expenditure on agricultural research and extension (R&E) during 1949-2016 and attempts to address the ad hoc model selection problem common in previous studies. Among the econometric modeling strategies in previous literature, Bayesian Model Averaging (BMA) and Bayesian Hierarchical Model (BHM) are two promising methods to solve the issue of model uncertainty. The rate-of-return estimates by BHM are preferable because BHM imposes fewer restrictions on lag structures and offers more reasonable lag shapes. By BHM, the internal rates of return (IRR) of Virginia's public expenditures on agricultural R&E are 26% and 42%, respectively. Nineteen percent of Virginia's agricultural productivity growth during 1949-2016 results from its R&E investments, and the

contribution of research to that growth is about twice of that of extension. One extra million dollar expenditure on research in 1992 would have brought a benefit of \$4.5 million, and the same expenditure in 1983 would have brought \$5.4 million in additional benefits. If the extra expenditure is spent on extension, it would have brought a benefit of \$6.1 million and \$6.3 million if the expenditure occurs in 1992 and 1983, respectively. Besides the modeling strategy, this study is distinguished from previous studies in that distributions of rates of return instead of only point estimates are obtained, which is missing in most studies.

Essays on Agricultural and Regional Development

Zhen Cheng

GENERAL AUDIENCE ABSTRACT

In a world of imbalance, food production and consumption exhibit great diversity among regions and countries. While farmers of developed economies benefit from up-to-date agricultural technology and produce more than they consume, households in the developing world are still facing food insecurity. This dissertation is composed of two manuscripts. Chapter 1 is about food consumption in developing countries. It analyzes household food demand behavior in the two West Africa countries – Niger and Nigeria with a focus on staple foods. Food demand behavior differs for rural and urban households and households of different income. Therefore, when evaluating the effects of policies and other impacts, policymakers and researchers should treat households with different attributes separately. Chapter 2 is on how to improve agricultural production within the context of a developed economy: it evaluates the returns to public expenditures on agricultural research and extension (R&E) in Virginia. Previous studies choose statistical models arbitrarily, and this study attempts to address this issue. It finds that Virginia's investments in agricultural R&E contribute to nineteen percent of the productivity growth in 1949-2016, and the contribution of research is about twice of that of extension.

Acknowledgments

The completion of my dissertation would not have been possible without the relentless support of my committee chair Dr. Jeffrey Alwang. I would also like to extend my deepest gratitude to the committee members: Dr. George Norton, for the unreserved imparting of his knowledge in agricultural economics that is helpful not only helpful for this dissertation but also for my lifelong academic career; Dr. Catherine Larochelle, for her enduring assistance and patience in the completion and writing of the work on food demand in Niger and Nigeria; Dr. Suqin Ge, for her thought-provoking questions that make me filling some important gaps in the studies.

I am also grateful to Dr. Klaus Moeltner and Dr. Aris Spanos. Dr. Moeltner's class ignited my interests in Bayesian econometrics that have greatly contributed to this dissertation. The guidance of Dr. Spanos reminds me the importance of statistical assumptions underlying econometric modeling, and his advice plays a significant role in the methodology of evaluating the returns to agricultural R&E.

List of Tables

List of Figures

Contents

Chapter 1

Demand Elasticities for Food Staples in Niger and Nigeria: A Three-Stage Approach

Food staples (cereals and other starchy foods) are the main source of calories for households in developing countries and play a major role in preventing hunger. The average annual per capita consumption of cereals is 201 kg in Niger (2011) and 138 kg in Nigeria (2013) while per capita consumption of tubers and roots averages 10.5 kg and 253 kg per year. Food staples account for 61% and 66% of Nigerien and Nigerian daily energy intake[1]. A supply shortage or a price surge can substantially aggravate food insecurity. During the 2008 food crisis, the global cereal price index rose to be 2.8 times higher than it was in 2000, and it remained high for several years. Globally, the crisis pushed approximately 130-155 million people into poverty and increased the prevalence of undernourishment by 6.8% in 2008 (United Nations, 2011). Understanding household behavior of staple food consumption, particularly in the form of demand elasticities, is essential for evaluating impacts of price changes on food demand, food security, and nutritional status.

In this study, household food demand systems are estimated for Niger and Nigeria to identify household food consumption behavior and fill knowledge gaps in staples demand. By mid-2017, the populations of Niger and Nigeria were 21.5 million and 190.9 million, respectively, and projected to reach 35.0 million and 264.1 million by 2030 (United Nations, 2017). Nigeria has a larger share of households living in urban areas than Niger (47.8% vs. 18.7% in 2015), and urbanization is predicted to reach 24.6% and 58.3% in 2030 (United Nations, 2018). By contrast, income growth has been erratic. Between 2010 and 2017, the GNI per capita growth rate fluctuated between -1.5% (2011) and 7.6% (2012) in Niger and between -3.7% (2016) and 5.3% (2014) in Nigeria. Rapid population growth, rising urbanization, and unstable income make food security an important issue in both countries. A GNI of only $387 per capita in 2017 (2010 constant U.S.

[1] Data are retrieved from FAO Food Security Statistics (http://www.fao.org/economic/ess/ess-fs/en/).

dollars) makes Niger one of the least developed countries in the world and more vulnerable to food insecurity.

Niger and Nigeria are similar not only in their names. They are neighbors and closely connected economically. Nigeria is the main regional trading partner of Niger. In 2000-04, 75-85% of Niger's millet and sorghum imports came from Northern Nigeria (Cornia and Deotti, 2008). Nigeria supplies almost 60% of Niger's grain deficit (Apaa-Okello, et al., 2015). Since grain production in Niger is volatile, the trade with Nigeria is an important safety net for Niger's food security. Nigeria is also an important market of Nigerien goods like livestock, so economic instability in Nigeria can lead to serious impacts on Niger's export earnings. When Nigeria suffered from low oil price in 2016, Niger's exports to Nigeria fell by 45.4% (International Monetary Fund, 2017). These connections make including the two countries in one study relevant.

Studies of household food demand in Africa are limited and subject to several shortcomings when available. After a review of food demand studies for Sub-Saharan African countries published between the mid-1980s and mid-1990s, Teklu (1996) warns against using these estimated elasticities in policy or impact analysis because they were usually not based on nationally representative data. In a meta-analysis on income elasticities of food demand in Africa, of the 2,028 elasticities from studies published in 1991-2015, only 116 are based on data collected in 2006-15 (Melo, et al., 2015). Another concern is that cross-price elasticities are rarely reported. Food substitution patterns can have significant implications for policy design. However, Cornelsen, et al. (2015) find that only slightly more than half of the 136 food demand studies published in 1990-2012 reported cross-price elasticities, and less than 5% of the cross-price elasticities were for African countries. Elasticities of individual staple food items are also infrequently estimated. By the meta-analysis by Melo, et al. (2015), only two of the 373 income elasticities of cereals are for sorghum. Unlike Asia, where rice is the dominant staple food across countries, staple consumption in Africa is diversified. Sorghum and millet are major staple foods for households in the Sahel, while tubers and roots are more important in humid areas of Nigeria, especially the southern coastal regions. Finally, yet importantly, we expect the food demand of households with different socio-economic status to respond differently to income and price changes. Average elasticities for the whole population do not allow evaluating the impact of food prices increase on calorie intakes of poor households, who are most vulnerable to food insecurity. However, the vast majority of studies did not distinguish between rural and urban households or households of different income groups (Melo, et al., 2015).

Methodological shortcomings such as ignoring censoring, using unit values instead of prices, failing to report unconditional elasticities and standard errors of elasticity estimates are also frequent in the literature on food demand in Africa. Censoring occurs when a household does not consume a food item, so the expenditure on the item is zero. Although ignoring censoring can potentially lead to inconsistent and biased estimates (Heien and Wessells, 1990), empirical studies frequently ignore the issue: fewer than 40% of the elasticities surveyed by Cornelsen, et al. (2015) addressed censoring. Applied demand literature relies on unit values (expenditure divided by quantity) to reflect prices and their variability. Unit values cannot directly substitute for market prices since unit values are subject to consumers' choice of quality and measurement errors (Deaton, 1988). Policy evaluations and impact analysis require unconditional elasticities, which are elasticities conditional on total expenditure. However, previous studies typically report elasticities conditioned on total food expenditure (e.g., Abdulai and Aubert, 2004; Bilgic and Yen, 2013), which are less helpful because food expenditure cannot be assumed to be constant. Further, many studies fail to report standard errors. Only 6% of the income elasticities surveyed by Melo, et al. (2015) and approximately half of the cross-price elasticities surveyed by Cornelsen, et al. (2015) are reported with standard errors. As pointed out by Cornelsen, it is difficult to know why studies do not report all estimation details.

We overcome these shortcomings by estimating a multi-stage censored demand system and adjusting for quality biases of unit values using regressions. Unconditional elasticities are calculated for each welfare quintile separately for rural and urban households to assess the responsiveness of food demand to price and income by household income. The approach of this study is similar to Boysen (2016) but differs in several aspects. The main difference is, instead of a two-stage system, we estimate a three-stage demand system, adopted by only a few studies (e.g., Ecker and Qaim, 2011; Edgerton et al., 1996; Jiang and Davis, 2007). The third stage allows estimation of elasticities of individual staple foods. The second methodological contribution consists of estimating the three stages simultaneously, which allows the direct computation of standard errors of elasticities from parameter estimates. The third difference is that we incorporate household characteristics into the demand system by demographic scaling so that households with the same income but different demographic characteristics can be distinguished. Boysen (2016) adopts the demographic translation approach, which does not allow the elasticities to vary with demographic characteristics.

Results confirm the disparity in household food demand behavior between rural/urban areas and between households of different income. The elasticities are in line with expectations in terms of their signs and the patterns when comparing rural/urban areas and welfare quintiles. However, the results are hard to interpret when we attempt to compare the two countries, resulting from the differences in the surveys and data processing between the countries. Beside elasticities, we also obtain the effects of demographic characteristics on household food consumption structure. Larger family size tends to worsen while better education tends to improve household nutrition status.

The rest of the article is structured as follows. We first introduce the data sets used and the procedures to obtain food price indices and expenditures. After that, we present details about model specification, estimation, and elasticities calculation. Descriptive statistics and elasticity estimates are presented and discussed. We discuss the findings and the proposed future work in the last section.

1.1 Data and data processing

Data sets were collected under the Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA) project led by the Development Research Group at the World Bank, in collaboration with national statistics institutes. For Niger, the National Survey of Household Living Conditions and Agriculture in 2011 included 3,859 households (2,343 rural and 1,516 urban). The analysis for Nigeria utilizes the second wave data of the General Household Survey-Panel (GHS-Panel) in 2012-13 that included 4,532 households (3,163 rural and 1,369 urban). Both surveys are nationally representative, representative of urban and rural areas, and include a household survey, a community survey, and an agriculture survey. Households were interviewed twice at post-planting and post-harvest time in order to capture the seasonality of agriculture and resulting consumption patterns. For instance, crop prices tend to be low during the post-harvest season because supply is abundant and high during the post-planting season because of shortage.

During the household survey, households were asked to recall the consumption of a wide variety of food items (125 in Niger and 105 in Nigeria) in the past seven days at each visit. The community survey of Niger includes a food price module that collected prices of 37 items from three sales spots. We use community data when available and household-specific unit values otherwise. The community survey of Nigeria includes not only fewer but also a different set of food items from the household survey. The food price modules of two visits are different in the kinds of food items and in whether varieties are differentiated. These issues make the food price information in the community survey unusable, and we only use household data to obtain food prices for Nigeria.

Aggregating food items into groups

A multi-stage demand system assumes households make consumption decisions in a stepwise process. Households first decide the allocation of their budget among major expenditure groups like food and non-food expenditures and then decide how to allocate their food budget among different items. It is necessary to aggregate food items into groups because estimating an equation for each item is impossible. Grouping is justified either by assuming that households' preferences satisfy separability conditions or by assuming external conditions like collinearity of prices in the Hicks-Leontief composite commodity theorem (Deaton and Muellbauer, 1980). Weak separability assumes the consumption of an item in one group does not influence the marginal rate of substitution between two items in another group, which is hard to test. The Hicks-Leontief composite commodity theorem states that several commodities can be treated as a single item when their prices move in parallel. The generalized composite commodity theorem by Lewbel (1996) relaxes the perfect collinearity assumption of prices and assumes that the ratio of the item price to group price index is independent of the group price index and income.

However, it is possible to group dis-similar commodities together based on these principles. We follow the convention in the food demand literature that aggregates food items based on their nutritional similarities. The first stage divides total household expenditure into a) *food-at-home* (FAH) and b) *food-away-from-home* (FAFH) *and all non-food expenditures*. FAFH is separated from FAH and we focus on FAH because of the lack of precise price information for FAFH. In the second stage, households are assumed to allocate their FAH expenditure across six groups: i) staple foods, ii) animal products, iii) vegetables and fruits, iv) legumes, nuts and seeds, v) oil, fat, sugar, spices, and other food complements, and vi) other FAH (e.g., beverages, snacks). The third stage decomposes staple food expenditure into six sub-groups: 1) millet, 2) sorghum, 3) rice, 4) corn, 5) other cereals and cereal-based foods (e.g., bread, biscuit, pasta), and 6) tubers and roots.

Food item prices

As the price of an item increases, a household reduces its consumption and/or relaxes quality requirements. Since products of lower prices and inferior quality replace those with higher prices and superior quality, the increase in unit values is less than that in market prices. As a result, substituting prices with unit values during estimation of demand relationships exaggerates households' responses to prices (Deaton, 1988). In surveys using clustering households in each community are surveyed within a short period, and it is reasonable to assume that all households in a community face the same price. As a result, observed differences in unit values are due to households' choice of quality, which in turns depends on household income and preferences (Deaton, 1997). Unit values can also be influenced by quantity since households that buy in large quantity usually receive lower prices.

Based on the assumptions, Deaton regresses the logarithmic unit value on the logarithmic total household expenditure and other characteristics to address the endogeneity of unit values. Spatial variations in prices can be substantially large in developing countries due to poor transportation infrastructure and market imperfections. Shin (2010) shows that market integration in Niger tends to be local and deteriorates rapidly during the 2004/05 food crisis. Therefore, community dummy variables are included as covariates, and their coefficients are taken as a proxy for market prices. Boysen (2016) extends Deaton's approach by adding dummy variables for measurement units and crop varieties.

Following Deaton and Boysen, for each food item k, we estimate the following model to obtain an adjusted price that is immune to the quality and quantity bias:

$$\begin{aligned} lnp_{ki} = b_{k0} + b_{k1}lnx_i + \boldsymbol{b}'_{kz}\boldsymbol{Z}_i + \boldsymbol{b}'_{ku}\boldsymbol{U}_{ki} + \boldsymbol{b}'_{kr}\boldsymbol{R}_i \\ +\boldsymbol{b}'_{ks}\boldsymbol{S}_{ki} + b_{k2}v_{2ki} + b_{kur}u_{ri} + \varepsilon_{ki} \end{aligned} \tag{1.1}$$

lnp_{ki} is the logarithmic unit value of item k reported by household i. lnx_i is the logarithm of household i's per capita real expenditure. $\boldsymbol{Z}_i$ is the vector of household socio-demographic variables: household size, number of household members by age groups except seniors, sex and marital status of household head, age and education of female household head or spouse of the male head[2]. $\boldsymbol{U}$ is the vector of dummy variables for the measurement units because

[2] The female household head or the spouse of male head is abbreviated as the primary female below.

households often reported quantities in local units (e.g., a heap of corn). $\boldsymbol{R}$ is the vector of dummy variables for communities, referred to as clusters below. $\boldsymbol{S}$ are dummy variables for sources of food (purchases, own-production, and in-kind payments or gifts), included because households might value food self-produced differently from purchased. $\boldsymbol{S}$ is not included for Nigeria since households only reported the value of purchased food. v_2 and u_r are binary variables that control for price differences between two visits and between rural and urban locations. The error term ε is assumed to be normal, independent, and identically distributed (NIID).

To obtain prices with quantity and quality bias adjusted, we suppose two representative households that only differ in locations (one from the reference cluster and the other from cluster r) and share identical characteristics on other aspects. By assumptions underlying Equation 1.1, the difference between unit values reported by the two households ($\hat{p}_{k0}$ and $\hat{p}_{kr}$) is only from the location:

$$ln\hat{p}_{kr} - ln\hat{p}_{k0} = \hat{b}_{kr} \tag{1.2}$$

$\hat{b}_{kr}$ is the estimated coefficient of the dummy variable of cluster r. Therefore,

$$\hat{p}_{kr}/\hat{p}_{k0} = \exp\left(\hat{b}_{kr}\right) \tag{1.3}$$

Equation 1.3 is the relationship between unit values in different locations after the quality and quantity bias has been adjusted, i.e., the relationship between market prices. $\hat{p}_{kr}$ could be obtained if we have the price in the reference cluster:

$$\hat{p}_{kr} = \hat{p}_{k0} \cdot \exp\left(\hat{b}_{kr}\right) \tag{1.4}$$

$\hat{p}_{k0}$ is estimated as the median of unit values reported by households in the first visit. The reference cluster is the cluster with the largest number of observations. We choose median instead of mean to avoid the influence of extreme values (Deaton, 1997).

To aggregate the expenditures of two visits, we convert the second-visit expenditure into its value in terms of the first-visit price by dividing it by a temporal deflator. The ratio of the second-visit price to that of the first visit $\hat{d}_k$ is estimated as:

$$\hat{d}_k \equiv \exp\left(\hat{b}_{k2}\right) \tag{1.5}$$

For Niger, we use community data to obtain prices for items included in the community survey. Though immune to the quality bias, prices from the community survey cannot be used directly because the community survey distinguished varieties for some food items while the household survey did not. We need to purge the community prices of the differences due to varieties. Similar to unit values from household data, community prices are also subject to the influence of local measurement units, the price difference between rural and urban areas, and price change between two visits. Therefore, a similar regression as that used to adjust household prices is applied to adjust the community prices:

$$lnp_k = a_{k0} + \boldsymbol{a}'_{kv}\boldsymbol{V}_k + \boldsymbol{a}'_{ku}\boldsymbol{U} + \boldsymbol{a}'_{kr}\boldsymbol{R} + a_{k2}v_2 + a_{kur}u_r + \varepsilon_k, \tag{1.6}$$

$\boldsymbol{V}$ is the vector of dummy variables for varieties of item k. $\boldsymbol{U}$, $\boldsymbol{R}$, v_2, and u_r are defined similarly as in Equation 1.1. The adjusted price and temporal deflator are derived similarly as Equations 1.4 and 1.5. Prices collected from different sale spots in a cluster are taken different observations.

Food group expenditures and price indices

To obtain the annual expenditure on each food item, we add up values of consumption from three sources and two visits and multiply it by 26, assuming each surveyed week represents half a year. Expenditures on food items are aggregated by food group and staple subgroup to obtain group and subgroup expenditures.

The price indices of food groups and staple subgroups are obtained using the Stone price index. Suppose $\overline{w}_{k|G}$ is the share of item k in the group or staple subgroup G averaged across households, the Stone price index of G for households living in cluster r is:

$$lnp_{Gr} = \sum_{k \in G} \overline{w}_{k|G} \ln \frac{\hat{p}_{kr}}{\bar{p}_k} \tag{1.7}$$

$\hat{p}_{kr}$ is the price of item k in cluster r, and $\bar{p}_k$ is the median of $\hat{p}_{kr}$ through clusters. Similarly, if $\overline{w}_{k|f}$ is the averaged share of item k in FAH, the price index of FAH lnp_{fr} is:

$$lnp_{fr} = \sum \overline{w}_{k|f} \ln \frac{\hat{p}_{kr}}{\bar{p}_k} \tag{1.8}$$

1.2 A three-stage demand system

A three-stage demand system accompanies the three-stage decomposition of household expenditure: a Working-Leser model for the first stage, a QUAIDS model for the second, and a censored QUAIDS for the third.

Stage 1: Working-Leser model

In the first stage, the share of FAH in total household expenditure (w_f) is modeled with a Working-Leser specification, assuming the shape of the Engel curve is determined by the log of food price and per capita aggregate expenditure. Since obtaining household total income in a developing-country context is difficult, we use total expenditure instead and use the elasticity of demand with respect to total expenditure as an approximation of the income elasticity. The original model is extended by incorporating household heterogeneity by demographic translation (Boysen, 2016):

$$w_f = \alpha_f + \gamma_f lnp_f + \beta_f \ln m + \lambda_f (\ln m)^2 + \boldsymbol{\tau}'\boldsymbol{z} + \varepsilon_f \tag{1.9}$$

lnp_f is the log of price index of FAH obtained by Equation 1.8. m is household per capita real expenditure. $\boldsymbol{z}$ is the vector of household demographic characteristics (the number of household members in each of the age groups 0-5, 6-15 and 16-60; the sex of household head; age and education of the primary female) and dummy variables for clusters. FAH expenditure elasticity η_f, uncompensated (Marshallian) price elasticity ε_f^U, and compensated (Hicksian) price elasticity ε_f^C, evaluated at $\overline{w}_f$ and $\overline{m}$, are derived as:

$$\eta_f = \frac{\beta_f + 2\lambda_f \ln \overline{m}}{\overline{w}_f} + 1 \tag{1.10}$$

$$\varepsilon_f^U = \frac{\gamma_f}{\overline{w}_f} - 1 \tag{1.11}$$

$$\varepsilon_f^C = \varepsilon_f^U + \overline{w}_f \eta_f = \frac{\gamma_f}{\overline{w}_f} + \beta_f + \overline{w}_f + 2\lambda_f \ln \overline{m} - 1 \tag{1.12}$$

Since the rural and urban subsamples are estimated separately, m is also normalized at sample median separately. Therefore, $\ln \overline{m}$ can be dropped from Equations 1.10-1.12 if the elasticities are evaluated at sample median ($\overline{m} = 1$). When calculating elasticities for each welfare quintile, $\overline{m}$ is substituted by mean of m of the quintile[3].

Stage 2: Quadratic Almost Ideal Demand System (QUAIDS)

A Quadratic Almost Ideal Demand System (Banks, Blundell and Lewbel, 1997) is used in the second stage to explain the demand for the six food groups. QUAIDS extends the Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980). Since most surveyed households consume all six groups (Table 1.1), the censoring adjustment is not made at this stage. Household socio-demographic characteristics are incorporated into the QUAIDS model by the demographic scaling method suggested by Ray (1983). Demographic scaling indicates the extra expenditure needed for a given household to achieve the same utility as that of a reference household (e.g., a married couple without children). It allows demographic variables to influence budget shares nonlinearly, so elasticities can vary with not only income and price but also with household characteristics. The final QUAIDS specification is:

$$w_{i|f} = \alpha_i + \sum_j \gamma_{ij}\, lnp_j + (\beta_i + \boldsymbol{\eta}_i'\boldsymbol{z}) \ln\left(\frac{x}{a(\boldsymbol{p})\overline{m}_0}\right) + \frac{\lambda_i}{b(\boldsymbol{p})}\left[\ln\left(\frac{x}{a(\boldsymbol{p})\overline{m}_0}\right)\right]^2 + \varepsilon_i \quad (1.13)$$

$$\ln a(\boldsymbol{p}) = \alpha_0 + \sum_i \alpha_i lnp_i + \sum_i \sum_j \gamma_{ij}\, lnp_i lnp_j \quad (1.14)$$

$$\overline{m}_0(\boldsymbol{z}) = 1 + \boldsymbol{\rho}'\boldsymbol{z} \quad (1.15)$$

$$b(\boldsymbol{p}) = \prod_j p_j^{\left(\beta_j + \boldsymbol{\eta}_j'\boldsymbol{z}\right)} \quad (1.16)$$

$w_{i|f}$ is the share of group i in FAH expenditure. lnp_i is the log price index of group i (Equation 1.7). x is household FAH expenditure normalized at sample median. The $\boldsymbol{z}$ vector is the same as in the first stage. To ensure adding-up condition ($\sum_i w_{i|f} = 1$) and homogeneity in prices and FAH expenditures $w_{i|f}(\lambda\boldsymbol{p}, \lambda x) = w_{i|f}(\boldsymbol{p}, x)$, restrictions 1.17, 1.18, and Slutsky symmetry restriction 1.19 are imposed:

$$\sum_i \alpha_i = 1, \sum_i \beta_i = 0, \sum_i \lambda_i = 0, \sum_k \gamma_{ik} = 0 \;\; \forall \;\; k, \sum_i \eta_{ir} = 0 \;\; \forall \;\; r \quad (1.17)$$

[3] Since the distribution of m has a very long tail, median is a better measurement of central tendency than mean. For each welfare quintile, mean and median are close and mean is used for calculating the elasticities.

$$\sum_{i} \gamma_{ik} = 0 \;\; \forall \;\; i \tag{1.18}$$

$$\gamma_{ik} = \gamma_{ki} \tag{1.19}$$

The expenditure elasticity of group i conditional on FAH expenditure is:

$$\eta_{i|f} = 1 + \frac{1}{w_{i|f}}\left[\beta_i + \boldsymbol{\eta}_{\boldsymbol{i}}'\boldsymbol{z} + \frac{2\lambda_i}{b(\boldsymbol{p})}\ln\left(\frac{x}{a(\boldsymbol{p})\overline{m}_0}\right)\right] \tag{1.20}$$

The uncompensated (Marshallian) price elasticity of group i with respect to the price of group j conditional on FAH expenditure is:

$$\begin{aligned}\varepsilon_{ij|f}^{U} = \frac{1}{w_{i|f}}\Bigg\{\gamma_{ij} - \left[\beta_i + \boldsymbol{\eta}_{\boldsymbol{i}}'\boldsymbol{z} + \frac{2\lambda_i}{b(\boldsymbol{p})}\ln\left(\frac{x}{a(\boldsymbol{p})\overline{m}_0}\right)\right] \times \left(\alpha_j + \sum_{l}\gamma_{jl}\ln p_l\right) \\ - \frac{(\beta_i + \boldsymbol{\eta}_{\boldsymbol{i}}'\boldsymbol{z})\lambda_i}{b(\boldsymbol{p})}\left[\ln\left(\frac{x}{a(\boldsymbol{p})\overline{m}_0}\right)\right]^2\Bigg\} - \Delta_{ij}\end{aligned} \tag{1.21}$$

Δ_{ij} equals one if $i = j$ and zero otherwise. Equation 1.21 represents the own-price elasticity if $i = j$ and the cross-price elasticity if $i \neq j$. Since we normalize prices at sample medians, elasticities evaluated at sample medians of prices ($\bar{p}_i = 1$) are:

$$\eta_{i|f} = 1 + \frac{1}{w_{i|f}}[\beta_i + \boldsymbol{\eta}_{\boldsymbol{i}}'\boldsymbol{z} + 2\lambda_i(\ln x - \ln \overline{m}_0)] \tag{1.22}$$

$$\begin{aligned}\varepsilon_{ij|f}^{U} = \frac{1}{w_{i|f}}\{\gamma_{ij} - \alpha_j[\beta_i + \boldsymbol{\eta}_{\boldsymbol{i}}'\boldsymbol{z} + 2\lambda_i(\ln x - \ln \overline{m}_0)] \\ - \lambda_i(\beta_i + \boldsymbol{\eta}_{\boldsymbol{i}}'\boldsymbol{z})(\ln x - \ln \overline{m}_0)^2\} - \Delta_{ij}\end{aligned} \tag{1.23}$$

Compensated (Hicksian) price elasticity conditional on FAH expenditure is obtained by the Slutsky decomposition:

$$\varepsilon_{ij|f}^{C} = \varepsilon_{ij|f}^{U} + w_j\eta_{i|f} \tag{1.24}$$

Stage 3: Censored QUAIDS

In the third stage, a censored QUAIDS model is estimated to explain household consumption of individual staple food items since censoring must be addressed. For instance, only 15.7% of households in urban Niger consume sorghum, and 35.4% of households in rural Nigeria consume millet. Ignoring censoring can lead to inconsistent and biased estimates (Heien and Wessells, 1990). To address censoring, Heien and Wessells (HW) propose a two-step procedure. The first step estimates probit regressions for each commodity to determine the probability that a household consumes the commodity; then inverse Mills ratios are computed for each household and applied to augment the share equations in the second step. HW show that the censored model greatly improves the goodness of fit and reduces both expenditure and own-price elasticities bias significantly for items consumed by only a few households.

Shonkwiler and Yen (1999) also estimate probit regressions in the first step but make improvements over HW's procedure in the second step. They assume that a latent variable $d^*_{ki} = \boldsymbol{\beta}'_k \boldsymbol{z}_{ki} + u_{ki}$, representing the net utility from consuming item k, determines the selection mechanism. $\boldsymbol{z}_{ki}$ is the vector of exogenous households variables, and the error term u_{ki} collects all the other factors influencing d^*_{ki}. Household i chooses to consume k if and only if it receives positive utility from consumption, i.e., the observed choice $d_{ki} = 1$ if and only if $d^*_{ki} > 0$. Given a consumption decision, the household decides how much to spend on k. The observed budget share $w_{ki} = d_{ki} \cdot w^*_{ki}$, where $w^*_{ki} = f(\boldsymbol{x}_{ki}, \boldsymbol{\gamma}_k) + \varepsilon_{ki}$ is the latent budget share. We adopt Shonkwiler and Yen's (SY) procedure because it produces consistent estimates, while estimates by HW's approach are inconsistent and the inconsistency becomes greater as more households do not consume an item.

SY suggest a two-step estimation procedure because the maximum likelihood estimation is difficult to implement for the correlation of error terms between equations. The first step estimates a probit model $\Pr(d_k = 1|\boldsymbol{z}_k) = \Phi(\boldsymbol{\beta}'_k \boldsymbol{z}_k)$ for each staple item k. For the model to be well identified, the independent variables in the binary selection equation ($\boldsymbol{z}_k$) must include at least one variable that is not in the share equation ($\boldsymbol{x}_k$). Otherwise, the coefficients will not have a structural interpretation. In addition to variables in the $\boldsymbol{z}$ vector in the first and second stages, $\boldsymbol{z}_k$ includes the share of food in household total expenditure and marital status of household heads. Since a household is unlikely to select a staple food independently of others, we estimate probit regressions of staple subgroups simultaneously and calculate the predicted probability $\widehat{\Pr}(d_k = 1|\boldsymbol{z}_k) = \Phi\left(\widehat{\boldsymbol{\beta}}'_k \boldsymbol{z}_k\right)$ and associated probability density $\phi\left(\widehat{\boldsymbol{\beta}}'_k \boldsymbol{z}_k\right)$. We only estimate probit regressions for staple items consumed by less than 95% of all households, otherwise $\widehat{\Phi}_k$ is assumed 0.9999 and $\widehat{\phi}_k$ takes the value of correponding normal density.

The second step augments the expenditure share equation in QUAIDS with $\hat{\Phi}_k$ and $\hat{\phi}_k$ to correct for censoring:

$$w_{k|1} = \widetilde{w}_{k|1} \cdot \hat{\Phi}_k + \delta_k \hat{\phi}_k + \varepsilon_k \tag{1.25}$$

$w_{k|1}$ is the expenditure share of subgroup k in staple foods, and $\widetilde{w}_{k|1}$ is the right-hand side of the share equation in the uncensored QUAIDS. Based on coefficient estimates, expenditure and uncompensated price elasticities of staple item k conditional on staple food expenditure, at sample price medians, are computed as (x_1 is expenditure on staple foods):

$$\eta_{k|1} = 1 + \frac{\overline{\Phi}_k}{w_{k|1}}[\beta_k + \boldsymbol{\eta}_k' \boldsymbol{z} + 2\lambda_k(\ln x_1 - \ln \overline{m}_0)] \tag{1.26}$$

$$\begin{aligned} \varepsilon_{kl|1}^U = \frac{\overline{\Phi}_k}{w_{k|1}}\{\gamma_{kl} - \alpha_l[\beta_k + \boldsymbol{\eta}_k' \boldsymbol{z} + 2\lambda_k(\ln x_1 - \ln \overline{m}_0)] \\ - \lambda_k(\beta_k + \boldsymbol{\eta}_k' \boldsymbol{z})(\ln x_1 - \ln \overline{m}_0)^2\} - \Delta_{kl} \end{aligned} \tag{1.27}$$

Estimation and calculation of unconditional elasticities

Since the uncensored QUAIDS is subject to the adding-up property, Equation 1.13 is omitted for Group 6 (other FAH) to avoid the singularity of the variance-covariance matrix. For the censored QUAIDS, the adding-up property does not hold, so all equations need to be estimated. Therefore, we cannot impose all the restrictions and omit the restriction on the λ parameters ($\sum_i \lambda_i = 0$). Finally, there are 12 equations to be estimated: one, five, and six equations for the first, second, and third stages, respectively.

The error terms are assumed to be correlated between equations but uncorrelated between observations, and all equations are estimated simultaneously by the nonlinear seemingly unrelated regression (NLSUR) method. We estimate the subsamples of rural and urban households separately for two reasons. First, urbanization could induce changes in food preferences. Second, urban households are more vulnerable to price fluctuations because they must rely on the market for the majority of food consumption while rural households produce part or all the foods for consumption. Verpoorten, et al. (2013) show that, during the 2008 food crisis, while self-reported food security improved in rural areas, urban households reported the worse situation in 2008 than in 2005 in 18 Sub-Saharan African countries.

After estimation, we calculate elasticities and their standard errors from the estimated parameters with the delta method. The unconditional expenditure and price elasticities of the six food groups of the second stage are computed by combining the first-stage elasticities and second-stage conditional elasticities, following Carpentier and Guyomard (2001):

$$\eta_i = \eta_{i|f}\eta_f \tag{1.28}$$

$$\varepsilon_{ij}^U = \varepsilon_{ij|f}^U + w_{j|f}\left(\frac{1}{\eta_{j|f}} + \varepsilon_f^U\right)\eta_{i|f}\eta_{j|f} + w_f w_{j|f}\eta_f\eta_{i|f}(\eta_{j|f} - 1) \tag{1.29}$$

$$\varepsilon_{ij}^C = \varepsilon_{ij|f}^C + w_{j|f}\varepsilon_f^C\eta_{i|f}\eta_{j|f} \tag{1.30}$$

Similarly, the unconditional elasticities of staple subgroups are computed as:

$$\eta_k = \eta_{k|1}\eta_1 \tag{1.31}$$

$$\varepsilon_{kl}^U = \varepsilon_{kl|1}^U + w_{l|1}\left(\frac{1}{\eta_{l|1}} + \varepsilon_{11}^U\right)\eta_{k|1}\eta_{l|1} + w_1 w_{l|1}\eta_1\eta_{l|1}(\eta_{l|1} - 1) \tag{1.32}$$

$$\varepsilon_{kl}^C = \varepsilon_{kl|1}^C + w_{l|1}\varepsilon_1^C\eta_{k|1}\eta_{l|1} \tag{1.33}$$

When calculating elasticities, household demographic variables take the value of the sample mean if they are continuous and sample mode if they are categorical. We allow demographic variables to differ between rural and urban areas and hold them constant across welfare quintiles, such that the influence of household expenditure on estimated elasticities could be observed.

1.3 Food consumption and household characteristics in Niger and Nigeria

Summary statistics of household food consumption and socio-demographic characteristics are presented in Tables 1.1 and 1.2, respectively. Summary statistics by welfare quintile are shown in Figures 1.1-1.5.

Household food consumption

In both countries, rural households allocate a larger share of their expenditure to FAH than urban households, and the food budget share generally falls from Q1 to Q5[4] and satisfies Engel's law. The budget share of FAFH is relatively small for Niger and rural Nigeria: <6% for each

[4] The first (poorest) to the fifth (wealthiest) welfare quintiles are abbreviated as Q1 to Q5.

quintile in Niger, <10% for each quintile in rural Nigeria. Therefore, ignoring FAFH due to the lack of price information is not expected to be a serious problem for Niger and rural Nigeria. However, ignoring FAFH is a potential problem for urban Nigeria. In urban Nigeria, the average share of FAFH is 10.9% and increases to 14.2% for Q5 (Figure 1.1).

The majority of the FAH budget is spent on staples and animal products: the two groups account for 68.4% of FAH budget in rural Niger, 65.8% in urban Niger, 61.3% in rural Nigeria, and 61.1% in urban Nigeria. Rural and poor households rely more heavily on staple foods. Staples account for 60.7% and 50.3% of FAH for households of Q1 in rural Niger and Nigeria, while for Q5 in urban areas the shares are only 28.2% and 32.8% (Figure 1.2). Consumption of animal products follows the opposite pattern: the budget share is higher among urban and better-off households. These patterns coincide with Bennett's law: households with lower income are heavily dependent on starchy-rich cereals and potatoes and tend to consume less well-balanced diets (Bennett, 1941). The budget shares of the other four groups are relatively constant when rural/urban areas or different welfare quintiles are compared. Nigerian households consume a more balanced diet than households in Niger: staples represent a lower share in Nigeria than Niger, compensated by larger shares of animal products, vegetables and fruits, and legumes, nuts, and seeds.

The composition of staple food consumption differs between countries and between rural/urban areas. Rural households in Niger spend more than half (54.1%) of their staple budget on millet, while in the urban areas the share of rice (35.1%) is more than twice that of millet (16.7%). In Nigeria, the largest share of staple budget is allocated to tubers and roots (38.2% in rural areas, 41.0% in urban areas), followed by rice (28.7% in rural areas, 38.9% in urban areas). The structure of consumption is consistent with agricultural production of the two countries: millet occupies almost one-third of cultivated land in Niger (Serra, 2015), while Nigeria produces 19% of cassava and 68% of yam worldwide (Sanginga and Mbabu, 2015).

Household socio-demographic characteristics

Households in the two countries share some common demographic characteristics. Urban households are smaller-sized, have fewer children, and are more likely to be headed by a female than rural households. Urban household heads and their spouses are significantly better educated than their rural counterparts. In Niger, only 31.5% of the primary females in rural areas have some level of education, while 45.5% and 20.9% in urban areas have primary and secondary education.

In Nigeria, only 18.3% of the rural primary females have experienced secondary education, compared with 48.0% in urban areas. In both rural and urban areas, household size and number of children decrease, and education of the primary female improves as income increases. Therefore, a large disparity in human development between rural/urban areas and between the rich and poor exists in the two countries. Differences are also apparent between countries. Compared with Niger, Nigerian households have smaller families and fewer children and are more likely to be headed by a female; household heads and their spouses are older and better educated. This is consistent with the fact that Nigeria is economically more developed than Niger, which is reflected in the food consumption patterns discussed above.

1.4 Results and discussion

In this section, we present and discuss the estimated elasticities (Tables 1.3 and 1.5-1.7[5], Figures 1.6-1.14) and the marginal effects of household characteristics on budget shares (Tables 1.8-1.11, Figures 1.15-1.22). Tables that include estimates by welfare quintile are shown in the appendices.

Model specification tests

The adjusted R^2 of the first-stage regression is over 0.9 in all four demand systems, indicating that models explain much of the variations in FAH budget shares across households. To test whether QUAIDS is a better fit for the data than AIDS, we test for the joint significance of the λ parameters using a Wald test. The λ parameters of the second and third stages are jointly significant at the 1% level in all four demand systems, supporting using the quadratic AIDS specification over a linear one. To test whether the censored QUAIDS specification is superior to the uncensored one, we perform a Wald test; the δ parameters are jointly significant at the 1% level in all four demand systems, and the null hypothesis that the δ parameters are all zero is rejected. Therefore, censoring needs to be addressed in the third stage. All socio-demographic variables and geographic dummy variables are jointly significant at the level of 1%. Most variables are significant at 1%, and all are significant at 10%.

[5] The significance of statistics in all tables of this article is represented as "***" for 1%, "**" for 5%, and "*" for 10%.

Expenditure and own-price elasticities of FAH

The expenditure and own-price elasticities of FAH are significant at 1%, and their signs are in line with expectations (positive for expenditure elasticities and negative for own-price elasticities) in all cases[6] (Table 1.3). FAH is a necessity (expenditure elasticity<1) except for the poorest households in Nigeria (Figure 1.6). The rural and poor households' demand for FAH is more responsive to income changes than households who are urban and rich, which is not surprising.

Since almost all the own-price elasticities are negative, we refer to the absolute values when discussing their magnitudes[7]. Given a demand system, the relationship between own-price elasticities and welfare quintiles can be understood by Equations 1.9 and 1.11. It is easy to prove that $\left|\varepsilon_f^U\right| < 1$ and decreases with welfare quintile if $\gamma_f > 0$, while $\left|\varepsilon_f^U\right| > 1$ and increases with quintile if $\gamma_f < 0$. Niger is the former case: the FAH budget share increases as price increases, so the FAH consumption drops less than the increase in price. Nigeria is consistent with the latter: the relationship between FAH share and price is negative; FAH consumption declines by a larger extent than the increase in price. In rural areas, the change in own-price elasticity between quintiles is trivial. From Q1 to Q5, $\left|\varepsilon_f^U\right|$ decreases from 0.868 to 0.811 in urban Niger while increases from 1.293 to 1.408 in urban Nigeria.

Timmer (1981) shows that compensated own-price elasticity of food falls with income in general, which is supported by many studies. However, the relationship is unclear for uncompensated elasticities. By Slutsky decomposition (Equation 1.24), as w_i and η_i decrease with income, the absolute values of uncompensated own-price elasticities are expected to decrease with income. As a result, the poor rural households are expected to be more price-responsive. This holds not only for food as a whole but also for particular food items like rice shown by Pinstrup-Andersen (1987). Therefore, it seems abnormal that food demand of Nigerian households are more price responsive than Nigerien households. It is also odd that in Nigeria, the urban and rich households are more own-price elastic than the rural and poor. A possible explanation is the larger FAFH budget in Nigeria, especially in urban areas. If households could easily shift from FAH to FAFH, the demand for FAH is more responsive to price.

[6] "All cases" means both countries, both rural and urban areas, and all welfare quintiles.
[7] Similarly, the vertical axis in all figures presenting own-price elasticities shows absolute values.

Expenditure elasticities of FAH groups

The expenditure elasticities of the six groups of FAH are positive and significant at 1% with only a single exception: the elasticity of legumes, nuts, and seeds is positive but insignificant for urban Q1 households in Niger (Table 1.3). The responsiveness of demand to income is more pronounced in rural areas compared to urban areas and decreases from Q1 to Q5 with a few exceptions (Figure 1.7)[8]. Therefore, similar to FAH as a whole, demand for food groups increases at a higher rate among poor households when income grows, which is consistent with expectations.

As expected, food staples are necessities. In Niger, the expenditure elasticity of staple foods is below animal products and vegetables and fruits but above the other groups, and only animal products are luxuries. Therefore, an increasing share of animal products and decreasing shares of other groups will be observed as income grows. In Nigeria, the expenditure elasticity of staples is always below those of other food groups. As a result, the staple budget increases slower than the other groups as income grows, so its share in FAH declines. It is not surprising to find that more groups are necessities in urban areas (the first four groups) than in rural areas (staples and animal products). Comparing the two countries, the demand for the two major groups (staples and animal products) is more elastic among Nigerien households than Nigerian households.

These results are generally rational, while some require further consideration. The results of Niger differ from the expectation that the expenditure elasticity of staples is lower than other groups. For evaluating the validity of the results, we collect unconditional expenditure and own-price elasticities of major good groups from several studies (Table 1.4). The elasticities from different studies are generally inconsistent, but commonly, the expenditure elasticities of staples (grain, cereals, tubers) are smaller than other groups. However, food staples in this study include processed food like bread and pasta, which are luxuries for the underdeveloped countries. This might be the reason why the expenditure elasticity of staples is not the lowest in Niger.

Own-price elasticities of FAH groups

Own-price elasticities of the six FAH groups are negative and significant at 1% with a single exception: that of other FAH is significant at 5% for Q1 in urban Niger (Table 1.3, Figure 1.8). The own-price elasticity of staples is less than one in Niger and rural Nigeria, while it is above

[8] Elasticities are larger in urban areas than in rural areas for vegetables and fruits in Niger and other FAH in Nigeria. In Niger, the elasticity of legumes, nuts, and seeds increases from Q1 to Q5 in urban areas.

but very close to one in urban Niger. Elasticities of other groups are mostly above or close to one. Therefore, staples are generally less own-price elastic than the other food groups. As prices of all groups increase, households try to preserve the consumption of staple foods, which coincides with its dominant role in the diet. Regular patterns do not exist when the own-price elasticities are compared between rural and urban areas. In Niger, demand for staples is slightly more price elastic in rural areas than in urban areas. In Nigeria, the own-price elasticity of staples in rural areas is significantly lower than that in urban areas.

In general, own-price elasticity of the food groups decreases from Q1 to Q5, but the variation is usually trivial. Only in the following cases, the changes with quintile are noticeable: staples in rural Nigeria (↓), animal products in Niger (↓), legumes, nuts, and seeds in urban Niger (↑), and other FAH in Niger and rural Nigeria (↓). The rising elasticity of legumes, nuts, and seeds with income in urban Niger seems abnormal at first. We should bear in mind the prediction by Timmer (1981) depends on the assumption that w_i and η_i fall with income. When the assumptions do not hold, it is not uncommon to find that poorest households have lower uncompensated own-price elasticities than middle-income households (Alderman, 1986) and own-price elasticity and income show an inverted-U shape relation (Anríquez, Daidone and Mane, 2013). In urban Niger, the budget share of legumes, nuts, seeds is constant with quintile, while expenditure elasticity increases dramatically through welfare quintiles. As a result, the own-price elasticity also increases.

Expenditure and own-price elasticities of the staple subgroups

The expenditure elasticities of staple subgroups are positive and significant at 1% with one exception: that of other cereals and cereal-based food is insignificant for Q1 in rural Nigeria (Table 1.5, Figure 1.9). The expenditure elasticities of staple subgroups generally decrease with welfare quintile, similar to staple foods as a whole. All staple subgroups are necessities in Nigeria and urban Niger, while subgroups 2-5 are luxuries in rural Niger. Therefore, as income grows in Niger, rural demand will increase by a larger extent that urban demand, not only for staples as a whole but for all staples except millet. In Nigeria, if the budget share of a staple item is larger in rural areas than in urban, the opposite is true for the expenditure elasticity. Since a large budget share is related to small elasticity, we can expect the composition of staple consumption in rural and urban areas converge as income grows.

Own-price elasticities of staple subgroups are generally negative and significant at 1% with a few exceptions, particularly corn in urban Nigeria (Table 1.5)[9]. The majority of the own-price elasticities are close to one. Demand for staple subgroups are more likely to be inelastic in Niger while elastic in Nigeria and tends to be more elastic in rural areas than in urban areas. The own-price elasticities are relatively constant across welfare quintiles, and large variations with quintile exist only in rural Nigeria: the elasticities of sorghum and corn increase while the that of rice decreases from Q1 to Q5.

Cross-price elasticities of staple subgroups

In rural Niger, when cross-price elasticities are significant, millet and tubers and roots are substitutes for other subgroups. The other subgroups are complements for each other except for the pair of rice and other cereals and cereal-based food (Table 1.6). Rural Nigerien households consume a large amount of millet, and they will switch from millet to other staples as millet price increases. In urban Niger, the majority of the significant cross-price elasticities are close to zero. Therefore, the demand for a staple subgroup is mainly influenced by its own price, and the substitution or complementation relationship between items is weak.

In Nigeria, when cross-price elasticities are significant, staple subgroups tend to be substitutes of each other in rural areas but complements in urban areas (Table 1.7). An explanation is that, for rural Nigerian households, the role of staples is to satisfy the demand for energy, so consuming one item is not quite different from consuming another. Urban households prefer a more diversified diet and tend to consume various items together.

In general, the cross-price elasticities do not vary greatly with welfare quintile (Figures 1.11-1.14). As the elasticity changes with quintile, it tends to move close to zero regardless of its sign. Therefore, the substitution and complementation relationships between staple items are weaker among the rich households than among the poor.

Marginal effects of household demographic characteristics on second-stage budget shares

The majority of the marginal effects of the demographic variables on budget shares are small in magnitude or insignificant, so we restrict our discussion to significant ones with relatively large magnitudes. A larger family has a significantly positive effect on the share of staple foods in

[9] Corn is ignored for urban Nigeria in Figure 1.10.

FAH, and its effects on the other groups are usually negative when significant (Tables 1.8 and 1.9). Therefore, larger families tend to consume more staples and less of other foods, leading to a more unbalanced diet and making the family's nutritional status worse. This result is consistent with the findings in several studies that assert an inverse relationship between family size and household food security (Olayemi, 2012). The effects are stronger for urban households than for rural households and slightly increases with welfare quintile (Figures 1.15-1.18). One extra adult in a household increases the share of staples in FAH by 2.2 and 1.7 percentage points in urban Niger and Nigeria, while the share increases by about 0.8 percentage points in rural areas.

The effects of the education of primary female on the budget share of staple foods are always negative when significant, and the effects are stronger in Niger than in Nigeria. In rural Niger, if the primary female has some level of education instead of being uneducated, the share of staples in FAH is 2.8 percentage points lower. The effects on other food groups are mainly positive. In urban Niger, secondary education has a particularly strong effect on the budget share of animal products. Compared with households whose primary female are uneducated and only received primary education, for households whose primary female received secondary education, the share of animal products in FAH is 4.3 and 3.7 percentage points higher. Therefore, education has the opposite effect of family size.

Marginal effects of household demographic characteristics on third-stage budget shares

In urban Niger and Nigeria, even when significant, the majority of the marginal effects of household characteristics on third-stage budget shares are less than one percentage point (Tables 1.10 and 1.11), which is universal through welfare quintiles (Figures 1.20 and 1.22). Therefore, household characteristics have a limited effect on the composition of staple consumption in urban areas. In rural Niger, a larger family is associated with a larger budget share of millet while smaller shares of other subgroups, while the effect of education of primary female is just the opposite: if the primary female has some education instead of being uneducated, the share of millet in staples is 5.7 percentage points lower. In rural Nigeria, a larger family is related to a smaller share of rice, while better education of the primary female is related to smaller shares of millet and tubers and roots and larger shares of other subgroups. Compared with illiteracy, the effects of primary education is trivial, and the effects of secondary education are relatively large.

1.5 Conclusion

This study contributes to the limited understanding of household food demand in Niger, one of the least developed countries in Africa, and Nigeria, the most populous country on the continent. Income growth is normally associated with a reduction in food insecurity and malnutrition, but it is not a guarantee. From 2008 to 2017, gross national income (GNI) per capita measured in 2010 US$ rose from 348 to 387 in Niger and from $1979 to $2354 in Nigeria. However, the prevalence of undernourishment increased from 12.4% to 14.4% in Niger and from 6% to 11.5% in Nigeria in 2008-2016[10].

The study employs national representative data and a rigorous econometric approach to estimate expenditure and price responsiveness of demand for aggregate FAH, food groups, and key staples. Similar data processing and estimation procedures are used to make the results comparable. Regression methods are used to purge the quality and quantity bias from unit values reported by households. The second and third stages of the demand system incorporates demographic variables by the method of demographic scaling that allows elasticities vary with demographics. The third stage addresses censoring using Shonkwiler and Yen's two-step method. A major contribution is that we calculate elasticities for households in each welfare quintile of rural and urban areas instead of treating the population as a whole. The results demonstrate how food demand behavior varies between rural and urban areas and between households of different economic status.

Staples are necessities for households in both countries, and staple shares in the diet are predicted to decrease as income grows. Results confirm the differences in staple food consumption behavior between rural and urban areas and among welfare groups. In comparison to urban and better-off households, poor rural households allocate a larger share of their food budget to staples, and their demand for staple foods as a whole and demand for each item is more responsive to income changes. Thus, poor rural households are more vulnerable to food insecurity when facing an income shock compared to their urban counterparts.

As food price increases, the decrease in staple consumption will be smaller than that of other foods, so households tend to preserve their staple consumption when prices rise. The relationship between own-price elasticities and welfare status is ambiguous. Within a country and

[10] Data are retrieved from https://data.worldbank.org/indicator/.

rural or urban areas, the own-price elasticity remains relatively stable through welfare quintiles for staples and its subgroups. In Niger, demand for staple foods is more responsive to a price change in rural areas than in urban areas, while it is the opposite in Nigeria.

In Nigeria, as income grows, the composition of staple consumption in rural and urban areas is predicted to converge, which is not observed for Niger. In rural Niger, millet is the substitute of other staple items except for tubers and roots, while other items are complements of each other. Substitution/complementarity between items is weak for urban Nigerien households. In Nigeria, staple items tend to be substitutes of each other in rural areas but complements in urban areas. The substitution/complementarity relationship tends to be weaker among the rich, so better-off households are more likely to maintain the composition of staple consumption when prices rise.

We also calculate the marginal effects of household demographics on food consumption. The two factors with the most significant effects are the family size and education of the primary female. While the former has a negative effect on households' nutrition status, better education of the primary female is related to a more balanced diet.

Some results from comparing the two countries are in line with expectations, but some are hard to interpret. Considering that Niger is less developed than Nigeria, it is not surprising to find that expenditure elasticities of staple foods and animal products are larger in Niger than in Nigeria. However, it is hard to explain that Nigerian households' demand for FAH is more responsive to income and price than that of Nigerien households. Actually, before estimating the elasticities, we observe some odd patterns in the data. The average share of foods[11] in total expenditure is larger in Nigeria than in Niger for both rural and urban areas (77.0% vs. 71.1% in rural areas, 66.8% vs. 53.9% in urban areas), which is abnormal considering that Nigeria is more developed than Niger. This is possibly the result of the differences in the surveys and data processing. For instance, Nigerien households were required to report the value of consumption from all sources, while Nigerian households only reported the value of purchased food. By the manuals accompanying the data sets, the aggregated household expenditure of Nigeria includes expenditures on education and health, while that of Niger does not include these. These factors undermine the comparability of the results between countries.

[11] We use the share of foods in total expenditure supplied with the data sets and split it into FAH and FAFH. Calculating the budget share of foods requires the calculation of household total expenditure, which needs a lot of work.

This study could be extended in several respects in future work. A key feature of food consumption in developing countries is that households, particularly those living in rural areas, produce a large fraction, or even all of the food they consume. Therefore, a rising food price brings extra profits to farmers, and the so-called *profit effect* can generate positive price elasticities of food in developing countries (Singh, et al., 1986). Evaluating the profit effect requires us to bring the production side into the analysis. Although the surveys used in this study include agricultural surveys that cover the production side, several serious problems make the data unusable. It is difficult to precisely split the reported inputs (e.g., land, labor, fertilizers, pesticides) among crops. Input prices are hard to obtain. Quantities of inputs are usually reported in local units, which are difficult to transform into metric units. These problems make the majority of the observations not usable, so we do not have enough observations to estimate production or profit functions for the production side.

Another extension could be incorporating FAFH in the analysis, considering the increasing importance of FAFH and processed food in the diet of households in developing countries. However, this requires additional information to be collected during the survey. For instance, the household survey of Nigeria ask the respondents to report the number of full meals (by breakfast/lunch/dinner) they had in the past week. We lack the information about what a meal contains (e.g., cereals, meat, vegetables). As a result, we are unable to decompose the expenditure on FAFH by the ingredients. Besides, part of the expenditure on meals at restaurants is service fees, making it improper to aggregate the expenditure on FAFH directly with that on FAH. For analyzing FAFH, it is more proper to ask a sample of households to make detailed bookkeeping of their diet in an extended period (e.g., a month). Asking the respondent to recall what they ate in a restaurant a week ago is difficult to receive accurate answers.

References

Abdulai, A., and D. Aubert. 2004. "A Cross-Section Analysis of Household Demand for Food and Nutrients in Tanzania." *Agricultural Economics* 31:67-79.

Alderman, H. 1986. "The Effect of Food Price and Income Changes on the Acquisition of Food by Low-Income Households." International Food Policy Research Institute.

Anríquez, G., S. Daidone, and E. Mane. 2013. "Rising Food Prices and Undernourishment: A Cross-Country Inquiry." *Food Policy* 38:190-202.

Apaa-Okello, J., M. Barry, C.A. Gueye, J. Jack, D. Marchettini, M. Nose, G.B. Li, and H. Wang. 2015. "Niger: Selected Issues." IMF Country Report No. 15/64. International Monetary Fund.

Banks, J., R. Blundell, and A. Lewbel. 1997. "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics* 79:527-539.

Bennett, M.K. 1941. "Wheat in National Diets." *Wheat Studies* 18:37-75.

Bilgic, A., and S.T. Yen. 2013. "Household Food Demand in Turkey: A Two-Step Demand System Approach." *Food Policy* 43:267-277.

Boysen, O. 2016. "Food Demand Characteristics in Uganda: Estimation and Policy Relevance." *South African Journal of Economics* 84:260-293.

Carpentier, A., and H. Guyomard. 2001. "Unconditional Elasticities in Two-Stage Demand Systems: An Approximate Solution." *American Journal of Agricultural Economics* 83:222-229.

Cornelsen, L., R. Green, R. Turner, A.D. Dangour, B. Shankar, M. Mazzocchi, and R.D. Smith. 2015. "What Happens to Patterns of Food Consumption when Food Prices Change? Evidence from A Systematic Review and Meta‐Analysis of Food Price Elasticities Globally." *Health Economics* 24:1548-1559.

Cornia, G.A., and L. Deotti. 2008. "Niger's 2005 Food Crisis: Extent, Causes and Nutritional Impact." EUDN Working Paper. European Development Research Network.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, MD: Johns Hopkins University Press.

---. 1988. "Quality, Quantity, and Spatial Variation of Price." *The American Economic Review* 78:418-430.

Deaton, A., and J. Muellbauer. 1980. "An Almost Ideal Demand System." *The American Economic Review* 70:312-326.

---. 1980. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

Ecker, O., and M. Qaim. 2011. "Analyzing Nutritional Impacts of Policies: An Empirical Study for Malawi." *World Development* 39:412-428.

Edgerton, D.L., B. Assarsson, A. Hummelmose, I.P. Laurila, K. Rickertsen, and P.H. Vale. 1996. *The Econometrics of Demand Systems: With Applications to Food Demand in the Nordic Countries*. Norwell, MA: Kluwer Academic Publishes.

Elijah Obayelu, A., V.O. Okoruwa, and O.I.Y. Ajani. 2009. "Cross-Sectional Analysis of Food Demand in the North Central, Nigeria: The Quadratic Almost Ideal Demand System (QUAIDS) Approach." *China Agricultural Economic Review* 1:173-193.

Fashogbon, A.E., and O.A. Oni. 2013. "Heterogeneity in Rural Household Food Demand and Its Determinants in Ondo State, Nigeria: An Application of Quadratic Almost Ideal Demand System." *Journal of Agricultural Science* 5:169-177.

Heien, D., and C.R. Wessells. 1990. "Demand Systems Estimation with Microdata: A Censored Regression Approach." *Journal of Business & Economic Statistics* 8:365-371.

International Monetary Fund. 2017. "Niger: Selected Issues." IMF Country Report No. 17/60.

Jiang, B., and J. Davis. 2007. "Household Food Demand in Rural China." *Applied Economic* 39:373-380.

Lewbel, A. 1996. "Aggregation Without Separability: A Generalized Composite Commodity Theorem." *The American Economic Review* 86:524-543.

Melo, P.C., Y. Abdul-Salam, D. Roberts, A. Gilbert, R. Matthews, L. Colen, S. Mary, and S.G.Y. Paloma. 2015. "Income Elasticities of Food Demand in Africa: A Meta-Analysis." JRC Technical Reports. Joint Research Centre, the European Commission.

Muhammad, A., J.L. Seale, B. Meade, and A. Regmi. 2011. "International Evidence on Food Consumption Patterns: An Update Using 2005 International Comparison Program Data." Technical Bulletin No. 1929. Economic Research Service, USDA.

Olayemi, A.O. 2012. "Effects of Family Size on Household Food Security in Osun State, Nigeria." *Asian Journal of Agriculture and Rural Development* 2:136-141.

Pinstrup-Andersen, P. 1987. "Macroeconomic Adjustment Policies and Human Nutrition: Available Evidence and Research Needs." *Food and Nutrition Bulletin* 9:1-19.

Ray, R. 1983. "Measuring the Costs of Children." *Journal of Public Economics* 22:89-102.

Sanginga, N., and A. Mbabu (2015) "Root and Tuber Crops (Cassava, Yam, Potato and Sweet Potato)." In *Proceedings of the An Action Plan for African Agricultural Transformation Conference.* Senegal, Dakar.

Serra, T. 2015. "Price Volatility in Niger Millet Markets." *Agricultural Economics* 46:489-502.

Shin, M. 2010. "A Geospatial Analysis of Market Integration: The Case of the 2004/5 Food Crisis in Niger." *Food Security* 2:261-269.

Shonkwiler, J.S., and S.T. Yen. 1999. "Two-Step Estimation of a Censored System of Equations." *American Journal of Agricultural Economics* 81:972-982.

Singh, I., L. Squire, J. Strauss, and B. World. 1986. *Agricultural Household Models: Extensions, Applications, and Policy*. Baltimore: Johns Hopkins University Press.

Teklu, T. 1996. "Food Demand Studies in Sub-Saharan Africa: A Survey of Empirical Evidence." *Food Policy* 21:479-496.

Timmer, C.P. 1981. "Is There 'Curvature' in the Slutsky Matrix?" *The Review of Economics and Statistics* 63:395-402.

United Nations. 2011. *The Global Social Crisis: Report on the World Social Situation 2011*. New York: United Nations Publications.

---. 2017. *World Population Prospects: The 2017 Revision*. New York: United Nations.

---. 2018. *World Urbanization Prospects: The 2018 Revision*. New York: United Nations.

Verpoorten, M., A. Arora, N. Stoop, and J. Swinnen. 2013. "Self-reported Food Insecurity in Africa During the Food Price Crisis." *Food Policy* 39:51-63.

Tables

Table 1.1. Summary of food consumption in Niger and Nigeria

	Rural				Urban			
	Budget shares (%)			% HHs consuming	Budget shares (%)			% HHs consuming
	Mean	Median	SE		Mean	Median	SE	
Niger								
FAH in total expenditure	66.3	67.8	11.7		49.5	50.4	13.0	
FAFH in total expenditure	4.8	1.8	7.1		4.4	2.5	5.8	
Stage 2: Share in FAH								
Staple foods	54.4	54.4	13.4	99.9	39.4	38.4	14.1	99.5
Animal products	14.0	12.0	11.0	95.3	26.4	24.3	15.5	98.9
Vegetables and fruits	6.7	6.2	3.8	99.1	10.0	9.6	5.2	99.5
Legumes, nuts and seeds	5.8	5.0	4.8	93.0	2.8	2.1	4.5	93.3
Oil, fat, sugar, spices, etc.	13.7	13.1	5.1	100.0	15.2	15.0	5.1	99.6
Other FAH	5.4	4.5	5.0	89.8	6.2	4.7	6.3	93.3
Stage 3: Share in staple foods								
Millet	54.1	54.1	20.8	99.5	16.7	12.8	14.9	83.0
Sorghum	9.6	5.8	12.1	48.6	1.7	0.0	5.2	15.7
Rice	11.4	8.6	12.7	71.1	35.1	33.9	16.6	97.1
Corn	10.8	0.0	14.7	47.4	14.7	12.9	13.3	72.5
Other cereals	7.6	4.5	9.1	67.0	20.4	17.6	15.8	93.3
Tubers and roots	6.4	4.6	7.2	72.9	11.0	9.0	9.7	92.1
Nigeria								
FAH in total expenditure	70.5	72.9	13.7		55.9	57.3	16.2	
FAFH in total expenditure	6.5	4.0	8.1		10.9	7.5	11.9	
Stage 2: Share in FAH								
Staple foods	41.7	41.0	15.6	100.0	39.0	37.9	14.3	99.9
Animal products	19.6	18.2	12.9	96.0	22.1	20.9	10.9	99.5
Vegetables and fruits	10.2	8.8	8.3	99.0	10.1	8.9	7.2	99.6
Legumes, nuts and seeds	11.9	8.8	12.0	89.6	9.0	7.6	7.4	93.1
Oil, fat, sugar, spices, etc.	12.6	8.3	12.5	99.8	11.4	7.7	11.8	99.2
Other FAH	4.0	0.7	8.9	56.2	8.5	5.2	11.3	80.6
Stage 3: Share in staple foods								
Millet	7.7	0.0	13.7	35.4	1.5	0.0	5.3	16.3
Sorghum	11.4	0.0	16.3	46.2	2.1	0.0	6.9	22.1
Rice	28.7	28.2	16.7	91.7	38.9	38.7	16.0	97.7
Corn	8.0	0.0	13.5	47.1	3.0	0.0	7.8	32.3
Other cereals	5.7	3.5	7.7	67.6	12.1	10.2	11.3	89.6
Tubers and roots	38.2	40.9	29.4	87.6	41.0	41.9	19.9	95.8

Table 1.2. Summary of household socio-demographic characteristics in Niger and Nigeria

	Rural			Urban		
	Continuous variables					
	Mean	Median	SE	Mean	Median	SE
Niger						
Number of HH members by age						
children, <=5	1.8	2.0	1.5	1.3	1.0	1.2
children, 6-15	1.9	2.0	1.8	1.8	1.0	1.9
adult	2.5	2.0	1.4	3.0	2.0	1.8
Age of primary female	35.2	32.0	13.9	37.5	35.0	14.1
Nigeria						
Number of HH members by age						
children, <=5	1.0	1.0	1.2	0.7	0.0	1.0
children, 6-15	1.9	2.0	1.7	1.3	1.0	1.5
adult	3.0	3.0	1.9	2.8	2.0	1.9
Age of primary female	43.2	40.0	15.2	44.9	43.0	14.8
	Categorical variables					
	% of HHs			% of HHs		
Niger						
HH headed by female			9.7			16.1
Education of primary female						
Some education			31.5			
Primary						45.5
>= Secondary						20.9
Nigeria						
HH headed by female			14.8			18.8
Education of primary female						
Primary			35.1			26.4
>= Secondary			18.3			48.0

Table 1.3. Expenditure and own-price elasticities, first and second stages

	Niger		Nigeria	
	Rural	Urban	Rural	Urban
Stage 1 FAH expenditure elasticity	0.876***	0.855***	0.958***	0.889***
	(0.011)	(0.010)	(0.008)	(0.016)
Stage 1 FAH uncompensated own-price elasticity	-0.955***	-0.848***	-1.078***	-1.341***
	(0.025)	(0.082)	(0.035)	(0.076)
Stage 2 unconditional expenditure elasticity				
Staple foods	0.770***	0.681***	0.662***	0.637***
	(0.018)	(0.016)	(0.026)	(0.026)
Animal products	1.448***	1.186***	0.943***	0.883***
	(0.047)	(0.033)	(0.038)	(0.040)
Vegetables and fruits	0.911***	0.986***	1.039***	0.876***
	(0.037)	(0.027)	(0.066)	(0.045)
Legumes, nuts, and seeds	0.664***	0.695***	1.311***	0.989***
	(0.039)	(0.168)	(0.076)	(0.050)
Oil, fat, sugar, spices, etc.	0.744***	0.672***	1.540***	1.382***
	(0.025)	(0.023)	(0.068)	(0.079)
Other FAH	0.978***	0.870***	1.024***	1.313***
	(0.068)	(0.070)	(0.160)	(0.123)
Stage 2 unconditional uncompensated own-price elasticity				
Staple foods	-0.931***	-0.891***	-0.548***	-1.008***
	(0.029)	(0.073)	(0.080)	(0.106)
Animal products	-1.394***	-0.989***	-1.136***	-1.136***
	(0.197)	(0.115)	(0.087)	(0.077)
Vegetables and fruits	-1.744***	-0.968***	-0.971***	-1.293***
	(0.185)	(0.104)	(0.068)	(0.165)
Legumes, nuts, and seeds	-0.982***	-4.628***	-0.881***	-0.846***
	(0.024)	(1.137)	(0.151)	(0.165)
Oil, fat, sugar, spices, etc.	-0.879***	-1.039***	-1.454***	-1.522***
	(0.107)	(0.119)	(0.059)	(0.109)
Other FAH	-1.616***	-1.528***	-1.313***	-1.167***
	(0.310)	(0.522)	(0.097)	(0.142)

Table 1.4. Unconditional expenditure and own-price elasticities from selected studies

	Elijah Obayelu, Okoruwa, and Ajani (2009)	Fashogbon and Oni (2013)	Muhammad, et al. (2011)		Cornelsen, et al. (2015)
	North Central, Nigeria	Ondo State, Nigeria	Nigeria	Low-income countries (range)	Low-income countries
Expenditure elasticity					
Grain/starch		0.74273			
Tubers	0.4078				
Cereals	1.1358		0.542	(0.314, 0.685)	
Animal products	1.4045	1.62913			
Meats			0.783	(0.707, 0.846)	
Fish			0.667	(0.571, 0.751)	
Dairy			0.81	(0.731, 0.877)	
Vegetables and fruits	1.104	0.79299	0.63	(0.512, 0.728)	
Legume	1.2817				
Fats and oils	0.0305	1.22936	0.557	(0.353, 0.691)	
Own-price elasticity					
Grain/starch		-0.75836			
Tubers	-0.385				
Cereals	-0.5312		-0.398	(-0.230, -0.502)	-0.61
Animal products	-0.0698	-1.15129			
Meats			-0.574	(-0.518, -0.621)	-0.78
Fish			-0.489	(-0.419, -0.551)	-0.8
Dairy			-0.594	(-0.536, -0.643)	-0.78
Vegetables and fruits	-0.0707	-1.16702	-0.462	(-0.375, -0.534)	-0.72
Legume	-0.3963				
Fats and oils	-0.4045	-1.86714	-0.408	(-0.259, -0.507)	-0.6

Table 1.5. Expenditure and own-price elasticities, the third stage

	Niger		Nigeria	
	Rural	Urban	Rural	Urban
Unconditional expenditure elasticity				
Millet	0.538*** (0.028)	0.681*** (0.016)	0.698*** (0.034)	0.808*** (0.060)
Sorghum	1.125*** (0.037)	0.674*** (0.016)	0.495*** (0.030)	0.665*** (0.057)
Rice	1.057*** (0.038)	0.680*** (0.016)	0.718*** (0.036)	0.708*** (0.032)
Corn	1.138*** (0.034)	0.681*** (0.016)	0.450*** (0.073)	0.584*** (0.056)
Other cereals and cereal-based food	1.167*** (0.044)	0.685*** (0.016)	0.551*** (0.116)	0.381*** (0.062)
Tubers and roots	0.900*** (0.104)	0.683*** (0.016)	0.791*** (0.035)	0.612*** (0.029)
Unconditional uncompensated own-price elasticity				
Millet	-0.908*** (0.028)	-0.976*** (0.013)	-1.122*** (0.088)	-0.973*** (0.237)
Sorghum	-1.178*** (0.104)	-1.014*** (0.002)	-1.099*** (0.093)	-1.050** (0.432)
Rice	-0.856*** (0.094)	-0.790*** (0.063)	-1.312*** (0.086)	-1.131*** (0.109)
Corn	-0.993*** (0.052)	-0.986*** (0.012)	-1.118*** (0.081)	0.053 (0.229)
Other cereals and cereal-based food	-1.234*** (0.051)	-0.954*** (0.016)	-0.688*** (0.091)	-0.686*** (0.119)
Tubers and roots	-1.570*** (0.277)	-0.507*** (0.180)	-1.177*** (0.074)	-1.028*** (0.099)

Table 1.6. Third-stage cross-price elasticities, Niger[12]

		Staple subgroups					
		1	2	3	4	5	6
Rural	1	-0.908***	0.103***	-0.020	0.162***	0.034***	-0.030
		(0.028)	(0.018)	(0.017)	(0.014)	(0.011)	(0.039)
	2	0.364***	-1.178***	-0.202**	-0.358***	-0.146***	0.142*
		(0.061)	(0.104)	(0.086)	(0.044)	(0.044)	(0.076)
	3	-0.027	-0.190**	-0.856***	-0.498***	0.048	0.227**
		(0.053)	(0.083)	(0.094)	(0.072)	(0.037)	(0.092)
	4	0.402***	-0.262***	-0.381***	-0.993***	-0.132***	-0.027
		(0.034)	(0.032)	(0.055)	(0.052)	(0.034)	(0.040)
	5	0.064	-0.153**	0.129**	-0.168**	-1.234***	-0.069
		(0.052)	(0.065)	(0.055)	(0.069)	(0.051)	(0.059)
	6	-0.182	0.306*	0.513**	-0.061	-0.111	-1.570***
		(0.277)	(0.158)	(0.203)	(0.117)	(0.089)	(0.277)
Urban	1	-0.976***	0.001	0.024	0.011	0.020	0.028***
		(0.013)	(0.002)	(0.025)	(0.012)	(0.016)	(0.010)
	2	0.017	-1.014***	0.011	0.026**	0.028*	0.050***
		(0.012)	(0.002)	(0.028)	(0.011)	(0.015)	(0.015)
	3	0.010	-0.008**	-0.790***	0.021*	0.031*	-0.153***
		(0.012)	(0.004)	(0.063)	(0.012)	(0.016)	(0.057)
	4	0.013	0.007***	0.047*	-0.986***	0.010	0.016
		(0.013)	(0.002)	(0.027)	(0.012)	(0.015)	(0.011)
	5	0.017	0.005***	0.048*	0.006	-0.954***	-0.020*
		(0.013)	(0.002)	(0.027)	(0.011)	(0.016)	(0.011)
	6	0.045***	0.033***	-0.450***	0.023	-0.038*	-0.507***
		(0.016)	(0.010)	(0.167)	(0.017)	(0.021)	(0.180)

[12] In Tables 1.6 and 1.7, the cross price elasticity is the demand for the subgroup on the left with respect to the price of the subgroup at the top. Subgroups 1-6 are millet, sorghum, rice, corn, other cereals and cereal-based food, tubers and roots.

Table 1.7. Third-stage cross-price elasticities, Nigeria

		Staple subgroups					
		1	2	3	4	5	6
Rural	1	-1.122*** (0.088)	0.065 (0.082)	-0.094 (0.078)	0.289*** (0.057)	-0.074** (0.034)	0.346*** (0.079)
	2	0.086 (0.073)	-1.099*** (0.093)	0.175*** (0.065)	0.074 (0.052)	0.013 (0.03)	0.333*** (0.069)
	3	-0.126** (0.052)	0.061 (0.05)	-1.312*** (0.086)	0.016 (0.035)	0.122*** (0.032)	0.633*** (0.073)
	4	0.392*** (0.064)	0.092 (0.071)	0.144** (0.062)	-1.118*** (0.081)	0.170*** (0.04)	-0.06 (0.071)
	5	-0.197** (0.093)	0.013 (0.09)	0.501*** (0.125)	0.347*** (0.092)	-0.688*** (0.091)	-0.441*** (0.127)
	6	0.096*** (0.034)	0.141*** (0.036)	0.429*** (0.053)	-0.068** (0.029)	-0.090*** (0.023)	-1.177*** (0.074)
Urban	1	-0.973*** (0.237)	0.414 (0.309)	0.169 (0.245)	-0.625** (0.243)	-0.457*** (0.169)	0.213 (0.232)
	2	0.329 (0.23)	-1.05** (0.432)	-0.447* (0.267)	-0.549* (0.286)	0.106 (0.196)	0.577** (0.254)
	3	0.114* (0.06)	-0.137 (0.085)	-1.131*** (0.109)	0.133** (0.055)	-0.115** (0.048)	0.033 (0.095)
	4	-0.539*** (0.205)	-0.613* (0.321)	0.473** (0.192)	0.053 (0.229)	0.494*** (0.135)	-0.778*** (0.191)
	5	-0.363*** (0.134)	0.096 (0.2)	-0.091 (0.139)	0.424*** (0.124)	-0.686*** (0.119)	0.026 (0.138)
	6	0.044 (0.056)	0.187** (0.083)	0.012 (0.093)	-0.226*** (0.055)	0.058 (0.047)	-1.028*** (0.099)

Table 1.8. Marginal effects of HH characteristics on second-stage budget shares, Niger (%)

	Staple foods	Animal products	Vegetables and fruits	Legumes, nuts, and seeds	Oil, fat, sugar, spices, etc.	Other FAH
Rural						
One more HH member						
child, <=5	0.752*** (0.241)	-0.344 (0.220)	-0.201*** (0.073)	-0.044 (0.073)	-0.120 (0.089)	-0.044 (0.088)
child, 6-15	1.476*** (0.229)	-0.903*** (0.221)	-0.277*** (0.059)	0.096 (0.070)	-0.056 (0.082)	-0.335*** (0.093)
adult	0.810*** (0.285)	-0.501* (0.258)	0.078 (0.111)	-0.180** (0.084)	-0.130 (0.099)	-0.077 (0.086)
Female HH head vs. male	-0.457 (1.286)	-1.125 (1.096)	1.016** (0.454)	-0.248 (0.414)	1.685*** (0.567)	-0.870 (0.535)
Primary female is five years older	0.480*** (0.148)	-0.316** (0.124)	-0.092* (0.048)	0.289*** (0.045)	-0.199*** (0.055)	-0.161*** (0.059)
Primary female education						
some vs. none	-2.759*** (0.708)	0.494 (0.630)	0.815*** (0.209)	0.290 (0.231)	0.417 (0.305)	0.743** (0.329)
Urban						
One more HH member						
child, <=5	1.186*** (0.411)	-0.903** (0.447)	-0.113 (0.093)	-0.160 (0.103)	0.149 (0.165)	-0.160 (0.132)
child, 6-15	2.181*** (0.255)	-1.711*** (0.292)	-0.301*** (0.084)	-0.366 (0.234)	0.350*** (0.116)	-0.153 (0.115)
adult	2.239*** (0.279)	-2.654*** (0.284)	-0.061 (0.078)	-0.320 (0.245)	0.706*** (0.124)	0.090 (0.109)
Female HH head vs. male	-2.537** (1.194)	1.880 (1.373)	0.415 (0.283)	1.551** (0.691)	-0.985** (0.493)	-0.324 (0.402)
Primary female is five years older	0.085 (0.160)	0.034 (0.190)	-0.131*** (0.039)	0.063 (0.046)	-0.016 (0.063)	-0.034 (0.073)
Primary female education						
primary vs. none	0.020 (0.693)	0.485 (0.776)	-0.311 (0.232)	-0.517* (0.285)	0.777*** (0.299)	-0.453 (0.409)
secondary vs. none	-1.573* (0.920)	1.255 (0.984)	0.603* (0.343)	-0.727 (0.483)	0.372 (0.406)	0.070 (0.629)
secondary vs. primary	-1.593* (0.822)	0.770 (0.896)	0.915*** (0.333)	-0.210 (0.295)	-0.405 (0.365)	0.523 (0.551)

Table 1.9. Marginal effects of HH characteristics on second-stage budget shares, Nigeria (%)

	Staple foods	Animal products	Vegetables and fruits	Legumes, nuts, and seeds	Oil, fat, sugar, spices, etc.	Other FAH
Rural						
One more HH member						
child, <=5	0.887***	-0.181	-0.247**	-0.340*	-0.411*	0.344**
	(0.296)	(0.134)	(0.098)	(0.202)	(0.234)	(0.164)
child, 6-15	1.135***	-0.151	-0.121	-0.310**	-0.590***	0.037
	(0.159)	(0.097)	(0.092)	(0.132)	(0.145)	(0.097)
adult	0.849***	-0.013	-0.033	-0.469***	-0.381**	0.047
	(0.147)	(0.086)	(0.083)	(0.108)	(0.168)	(0.095)
Female HH head vs. male	-0.470	-0.349	-0.527	0.097	1.145	0.103
	(0.477)	(0.491)	(0.420)	(0.482)	(0.799)	(0.532)
Primary female is five years older	0.224***	-0.083	0.040	-0.122*	-0.095	0.037
	(0.085)	(0.057)	(0.060)	(0.062)	(0.074)	(0.053)
Primary female education						
primary vs. none	-0.086	-1.137***	0.092	0.312	0.174	0.657**
	(0.455)	(0.358)	(0.297)	(0.482)	(0.463)	(0.318)
secondary vs. none	-1.027**	-0.128	-0.711**	-0.329	0.587	1.609***
	(0.475)	(0.438)	(0.295)	(0.512)	(0.637)	(0.496)
secondary vs. primary	-0.941*	1.009**	-0.804**	-0.647	0.413	1.017**
	(0.487)	(0.410)	(0.318)	(0.605)	(0.632)	(0.512)
Urban						
One more HH member						
child, <=5	3.364***	0.401	-0.310	-0.236	-1.525**	-1.694***
	(0.839)	(0.400)	(0.204)	(0.243)	(0.761)	(0.558)
child, 6-15	2.415***	-0.333	-0.124	0.310*	-0.956**	-1.312***
	(0.393)	(0.245)	(0.155)	(0.181)	(0.385)	(0.308)
adult	1.747***	0.019	-0.179	0.049	-0.944***	-0.691***
	(0.343)	(0.203)	(0.110)	(0.139)	(0.315)	(0.265)
Female HH head vs. male	0.715	-1.127	-0.350	0.361	1.143	-0.743
	(1.311)	(0.857)	(0.437)	(0.544)	(1.105)	(0.874)
Primary female is five years older	0.552***	0.176	0.051	0.075	-0.218	-0.636***
	(0.210)	(0.137)	(0.081)	(0.084)	(0.160)	(0.164)
Primary female education						
primary vs. none	1.158	0.562	0.012	0.026	-0.659	-1.099
	(1.520)	(0.752)	(0.412)	(0.566)	(1.250)	(1.034)
secondary vs. none	-0.828	4.289***	-0.148	-0.885	-0.953	-1.475
	(1.421)	(0.841)	(0.444)	(0.560)	(1.064)	(1.067)
secondary vs. primary	-1.985	3.728***	-0.161	-0.912*	-0.294	-0.376
	(1.418)	(0.758)	(0.384)	(0.474)	(1.122)	(0.925)

Table 1.10. Marginal effects of HH characteristics on third-stage budget shares, Niger (%)

	Millet	Sorghum	Rice	Corn	Other cereals and cereal-based food	Tubers and roots
Rural						
One more HH member						
child, <=5	1.515*** (0.194)	-0.312*** (0.103)	-0.447*** (0.107)	-0.552*** (0.061)	-0.561*** (0.093)	-0.008 (0.138)
child, 6-15	1.862*** (0.198)	-0.487*** (0.094)	-0.432*** (0.111)	-0.655*** (0.062)	-0.392*** (0.088)	-0.292 (0.178)
adult	1.364*** (0.223)	-0.189 (0.122)	-0.337** (0.157)	-0.585*** (0.069)	-0.550*** (0.113)	-0.354** (0.141)
Female HH head vs. male	-1.662** (0.747)	-2.063*** (0.500)	1.717*** (0.652)	-0.492* (0.285)	1.232** (0.532)	1.344** (0.658)
Primary female is five years older	0.792*** (0.125)	-0.154*** (0.054)	-0.172** (0.067)	-0.229*** (0.039)	-0.225*** (0.056)	-0.026 (0.080)
Primary female education						
some vs. none	-5.704*** (0.706)	0.718** (0.285)	1.426*** (0.413)	0.180 (0.180)	-0.308 (0.310)	1.283*** (0.374)
Urban						
One more HH member						
child, <=5	-0.032** (0.014)	0.009*** (0.001)	-0.045** (0.023)	-0.004 (0.007)	-0.010 (0.014)	0.050** (0.021)
child, 6-15	-0.021* (0.011)	0.001 (0.001)	0.020 (0.017)	0.022*** (0.007)	-0.053*** (0.012)	0.027** (0.014)
adult	0.029*** (0.011)	-0.003*** (0.001)	-0.206*** (0.024)	0.017 (0.011)	0.138*** (0.018)	0.014 (0.017)
Female HH head vs. male	0.173*** (0.046)	-0.022*** (0.004)	-0.800*** (0.089)	-0.029 (0.037)	0.489*** (0.070)	0.198*** (0.061)
Primary female is five years older	-0.043*** (0.008)	0.007*** (0.001)	0.129*** (0.016)	0.003 (0.007)	-0.096*** (0.014)	-0.011 (0.011)
Primary female education						
primary vs. none	0.229*** (0.044)	-0.027*** (0.002)	-0.660*** (0.073)	-0.075*** (0.023)	0.512*** (0.057)	0.094* (0.051)
secondary vs. none	0.324*** (0.060)	-0.043*** (0.004)	-1.317*** (0.135)	-0.101*** (0.034)	1.036*** (0.125)	0.179* (0.097)
secondary vs. primary	0.095** (0.038)	-0.015*** (0.002)	-0.656*** (0.100)	-0.025 (0.023)	0.523*** (0.107)	0.085 (0.074)

Table 1.11. Marginal effects of HH characteristics on third-stage budget shares, Nigeria (%)

	Millet	Sorghum	Rice	Corn	Other cereals and cereal-based food	Tubers and roots
			Rural			
One more HH member						
child, <=5	-0.088	0.113	-1.314***	0.454***	0.053	0.282
	(0.087)	(0.130)	(0.233)	(0.113)	(0.139)	(0.292)
child, 6-15	-0.083	0.476***	-0.683***	0.080	0.111	-0.346
	(0.053)	(0.086)	(0.157)	(0.093)	(0.101)	(0.215)
adult	-0.052	-0.008	-0.683***	0.404***	0.064	-0.085
	(0.054)	(0.094)	(0.202)	(0.085)	(0.096)	(0.212)
Female HH head vs. male	-0.856	-0.895	5.724***	1.377**	-1.669**	-2.111
	(0.598)	(0.856)	(1.820)	(0.683)	(0.710)	(1.775)
Primary female is five years older	-0.163***	0.211***	-0.099	0.081	-0.353***	0.378**
	(0.051)	(0.072)	(0.155)	(0.067)	(0.071)	(0.164)
Primary female education						
primary vs. none	-0.248	-0.683**	0.104	0.965***	0.508	-0.719
	(0.213)	(0.309)	(0.602)	(0.328)	(0.387)	(0.654)
secondary vs. none	-2.678***	1.043**	4.969***	2.729***	0.967*	-6.462***
	(0.324)	(0.522)	(1.020)	(0.572)	(0.536)	(1.144)
secondary vs. primary	-2.430***	1.652***	4.865***	1.764***	0.458	-5.802***
	(0.311)	(0.482)	(1.009)	(0.630)	(0.581)	(1.161)
			Urban			
One more HH member						
child, <=5	0.043	0.028	0.142	-0.295**	-1.127***	1.822***
	(0.060)	(0.073)	(0.484)	(0.130)	(0.432)	(0.554)
child, 6-15	-0.009	-0.058	-0.208	-0.024	0.536**	0.153
	(0.036)	(0.044)	(0.323)	(0.073)	(0.216)	(0.313)
adult	0.025	0.015	0.212	-0.161***	-0.029	0.204
	(0.029)	(0.033)	(0.285)	(0.061)	(0.164)	(0.285)
Female HH head vs. male	-0.275	0.167	2.014	0.375	-0.844	-1.750
	(0.259)	(0.291)	(1.594)	(0.444)	(1.074)	(1.670)
Primary female is five years older	0.062*	-0.011	-0.335	-0.114*	-0.434**	0.914***
	(0.032)	(0.038)	(0.229)	(0.059)	(0.181)	(0.225)
Primary female education						
primary vs. none	0.144	0.420**	0.530	-0.440	-1.147	-1.189
	(0.134)	(0.182)	(1.249)	(0.305)	(0.887)	(1.295)
secondary vs. none	-0.163	1.008**	1.194	-0.896*	0.582	-2.911*
	(0.280)	(0.394)	(1.800)	(0.525)	(1.308)	(1.626)
secondary vs. primary	-0.307	0.494	0.663	-0.456	1.728*	-1.721
	(0.277)	(0.333)	(1.553)	(0.498)	(1.000)	(1.478)

Figures

Figure 1.1. First-stage mean budget shares by welfare quintile (%)

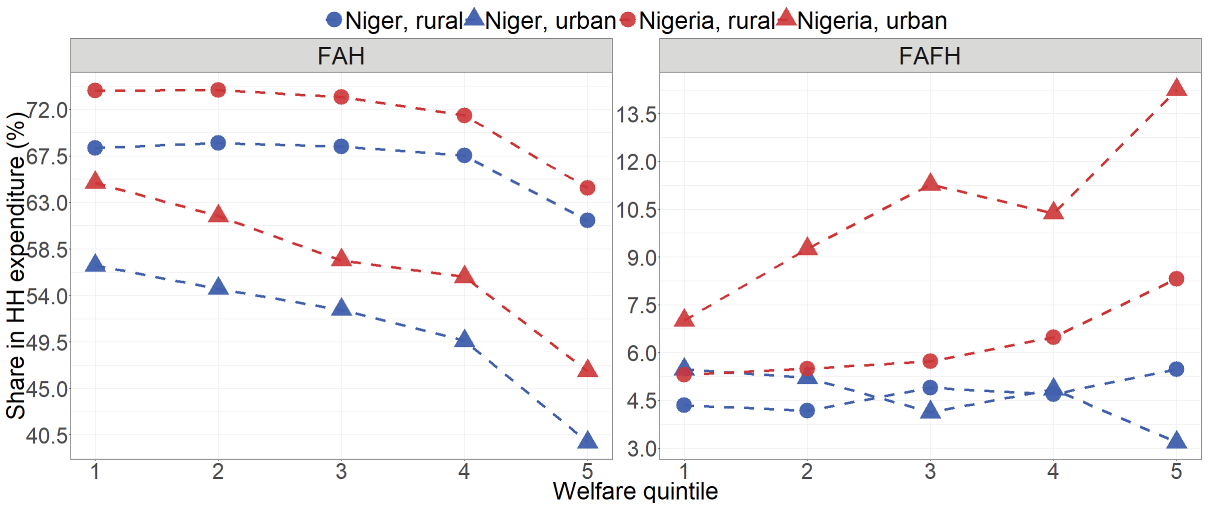


Figure 1.2. Second-stage mean budget shares by welfare quintile (%)

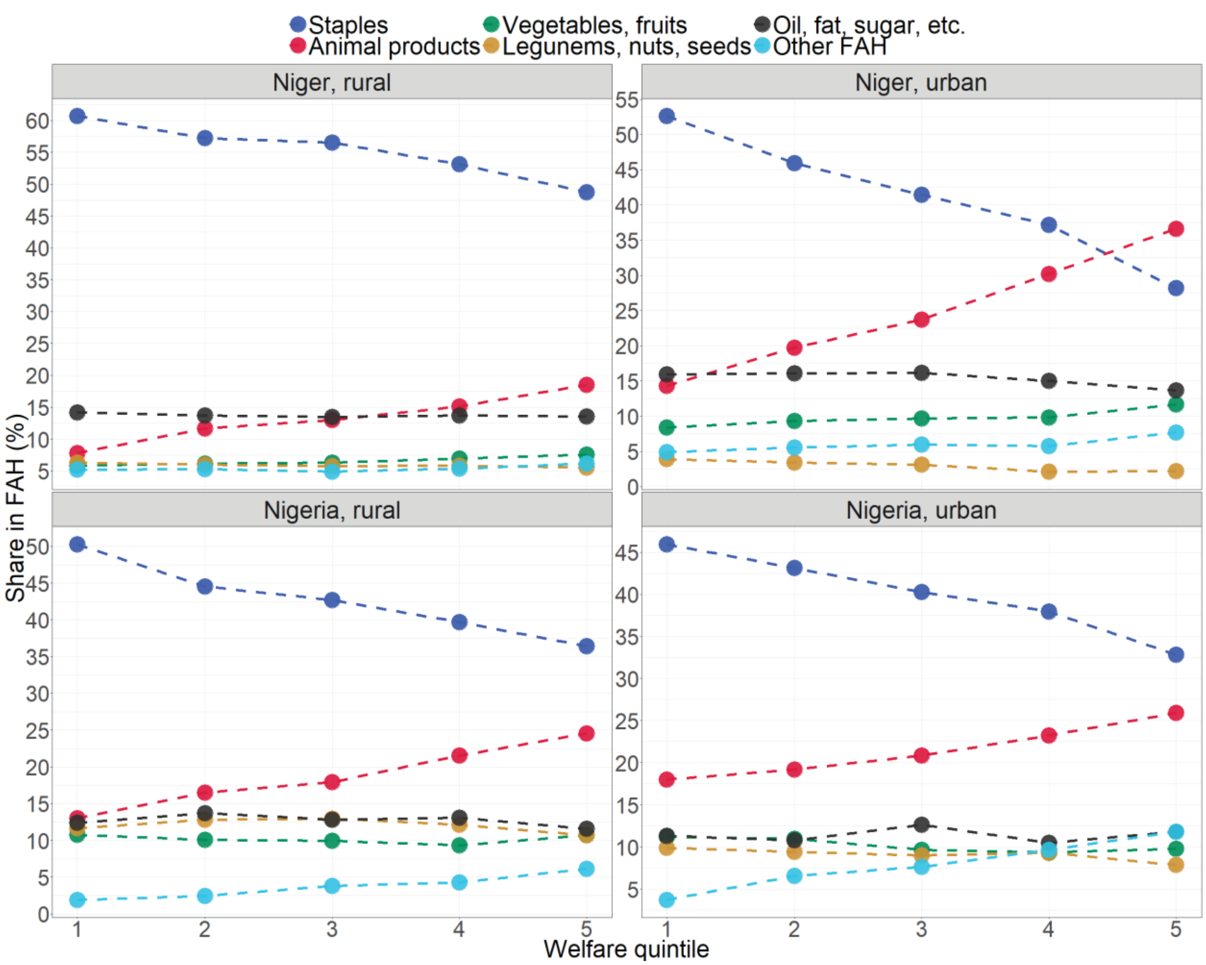


Figure 1.3. Third-stage mean budget shares by welfare quintile (%)

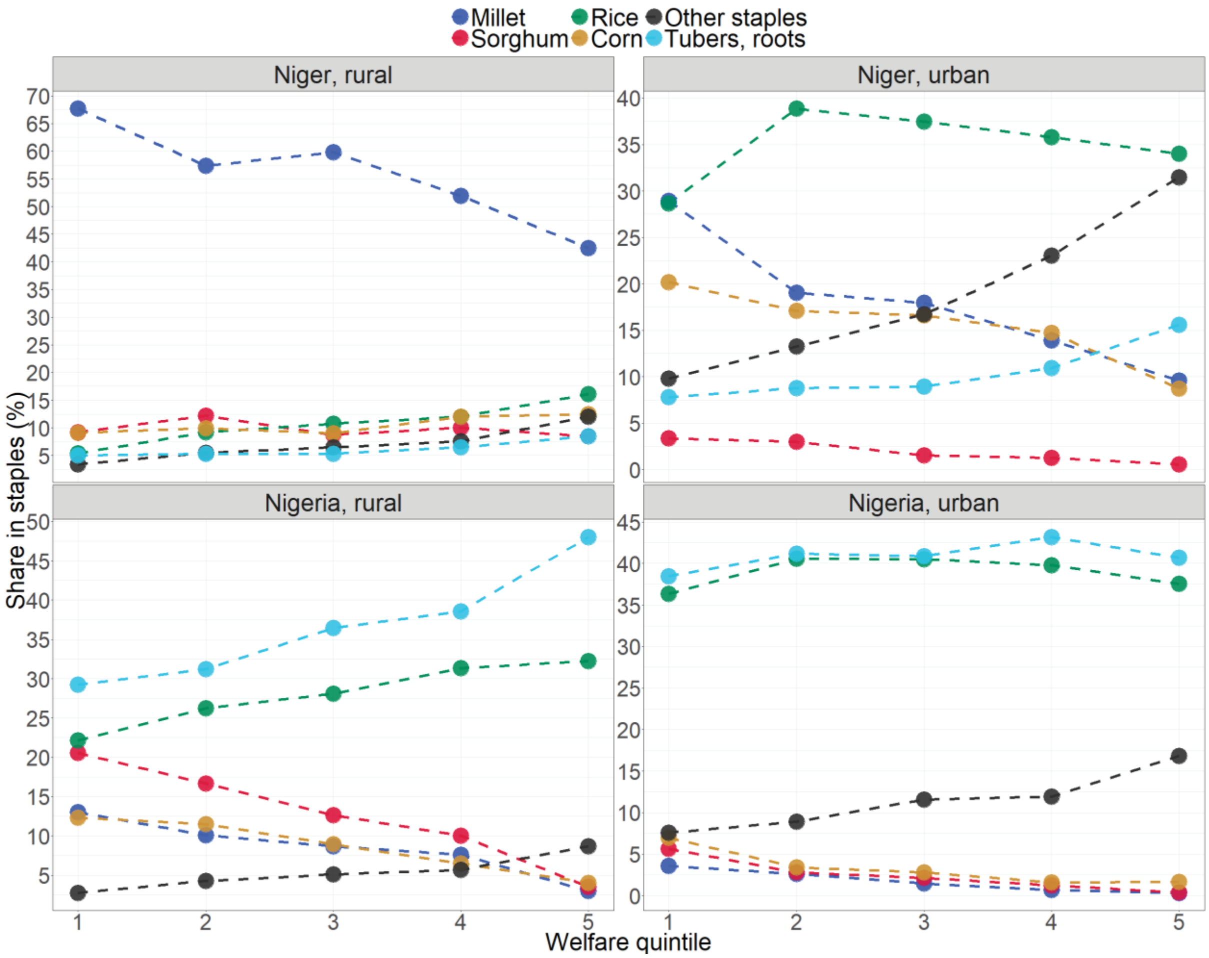


Figure 1.4. Household characteristics: mean of continuous variables by welfare quintile

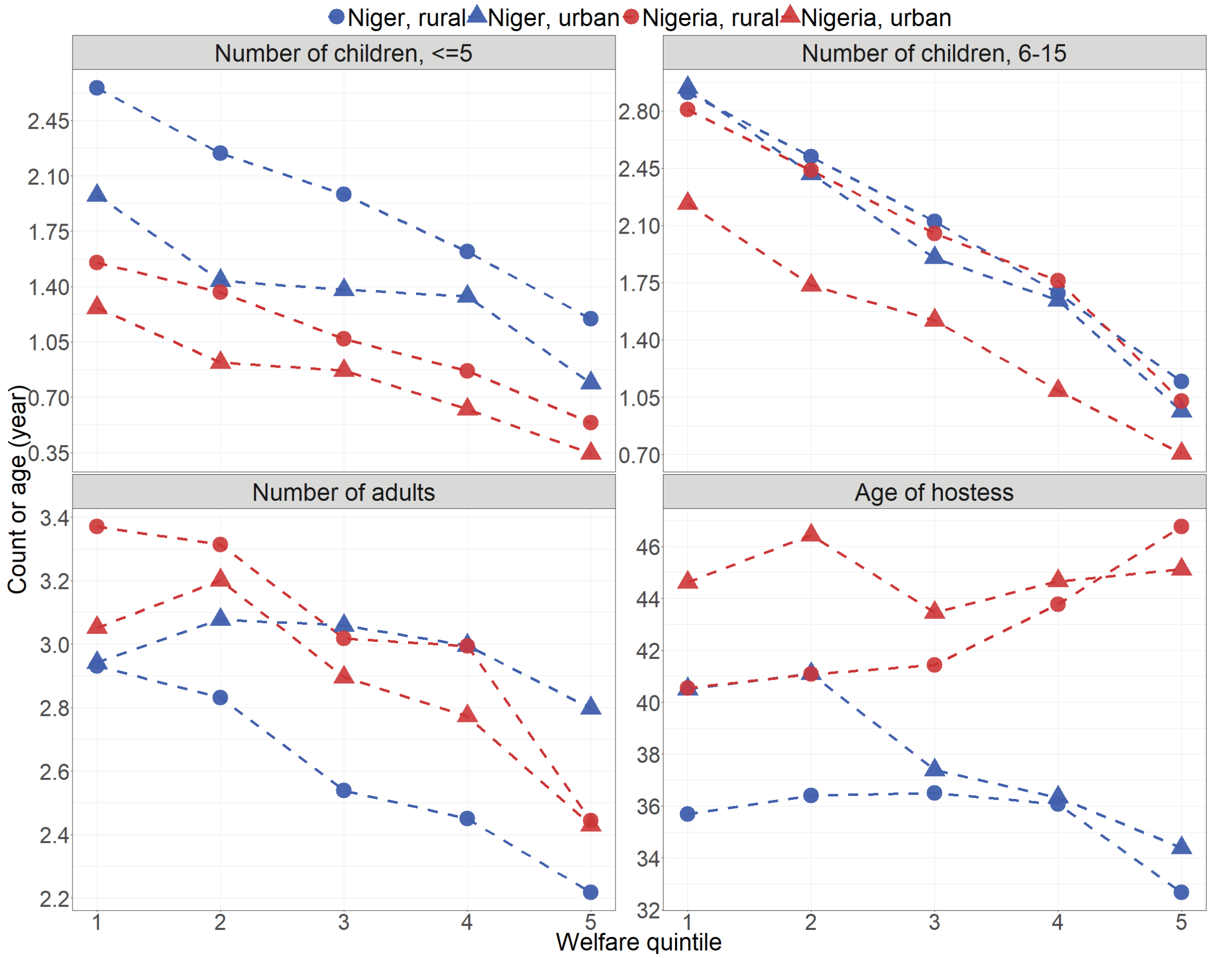


Figure 1.5. Household characteristics: categorical variables by welfare quintile

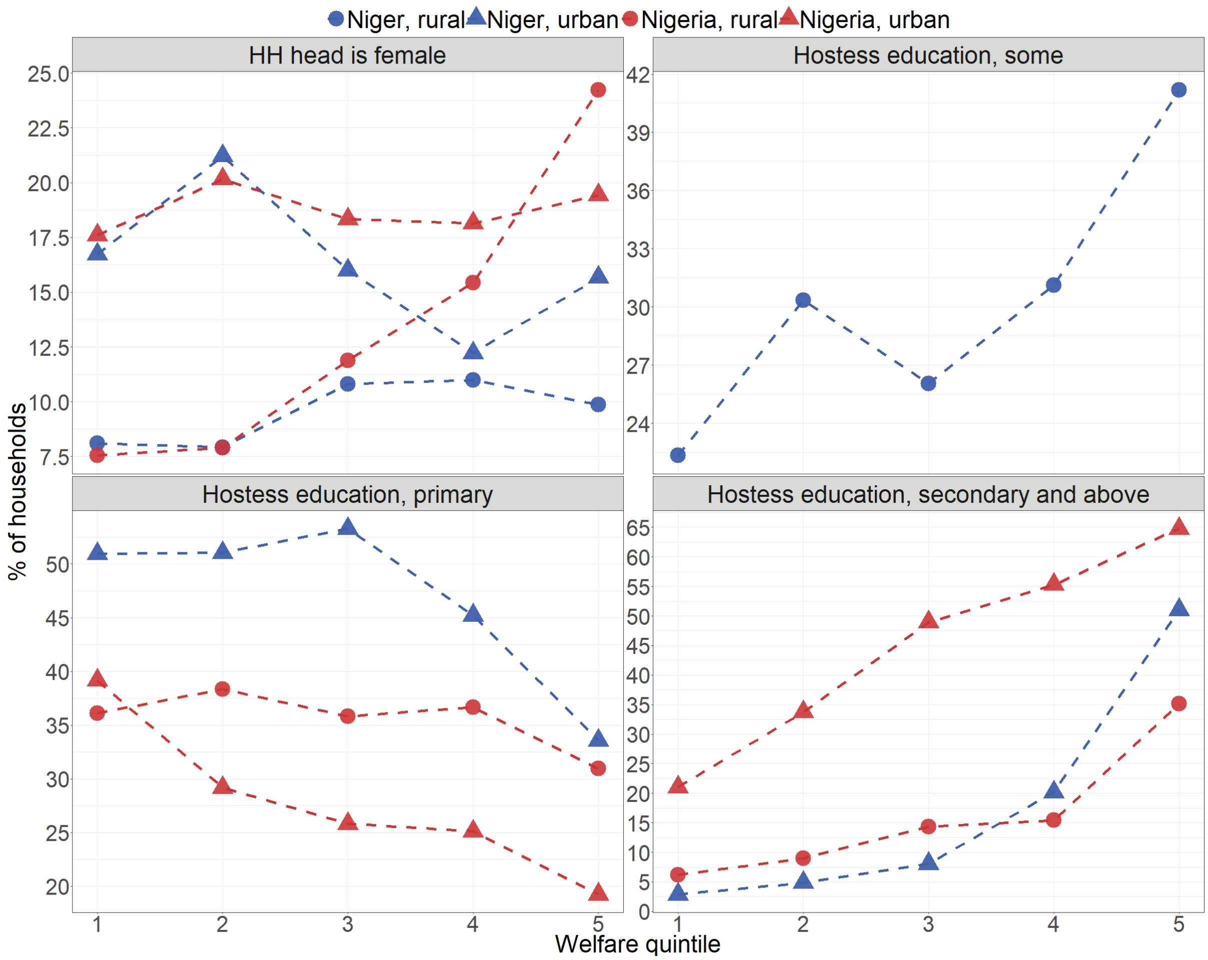


Figure 1.6. First-stage expenditure and own-price elasticities by welfare quintile

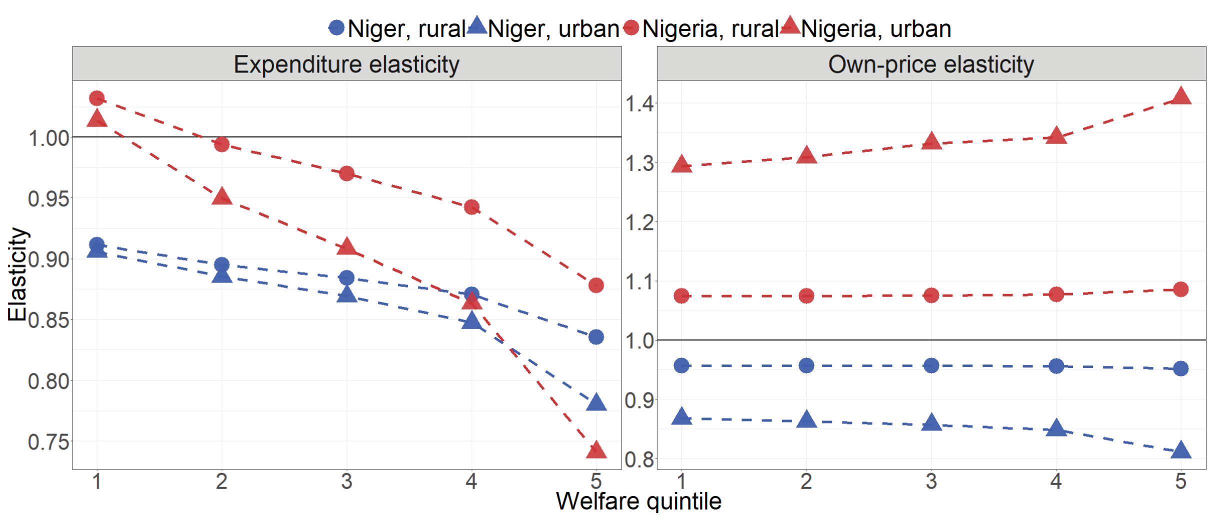


Figure 1.7. Second-stage expenditure elasticities by welfare quintile

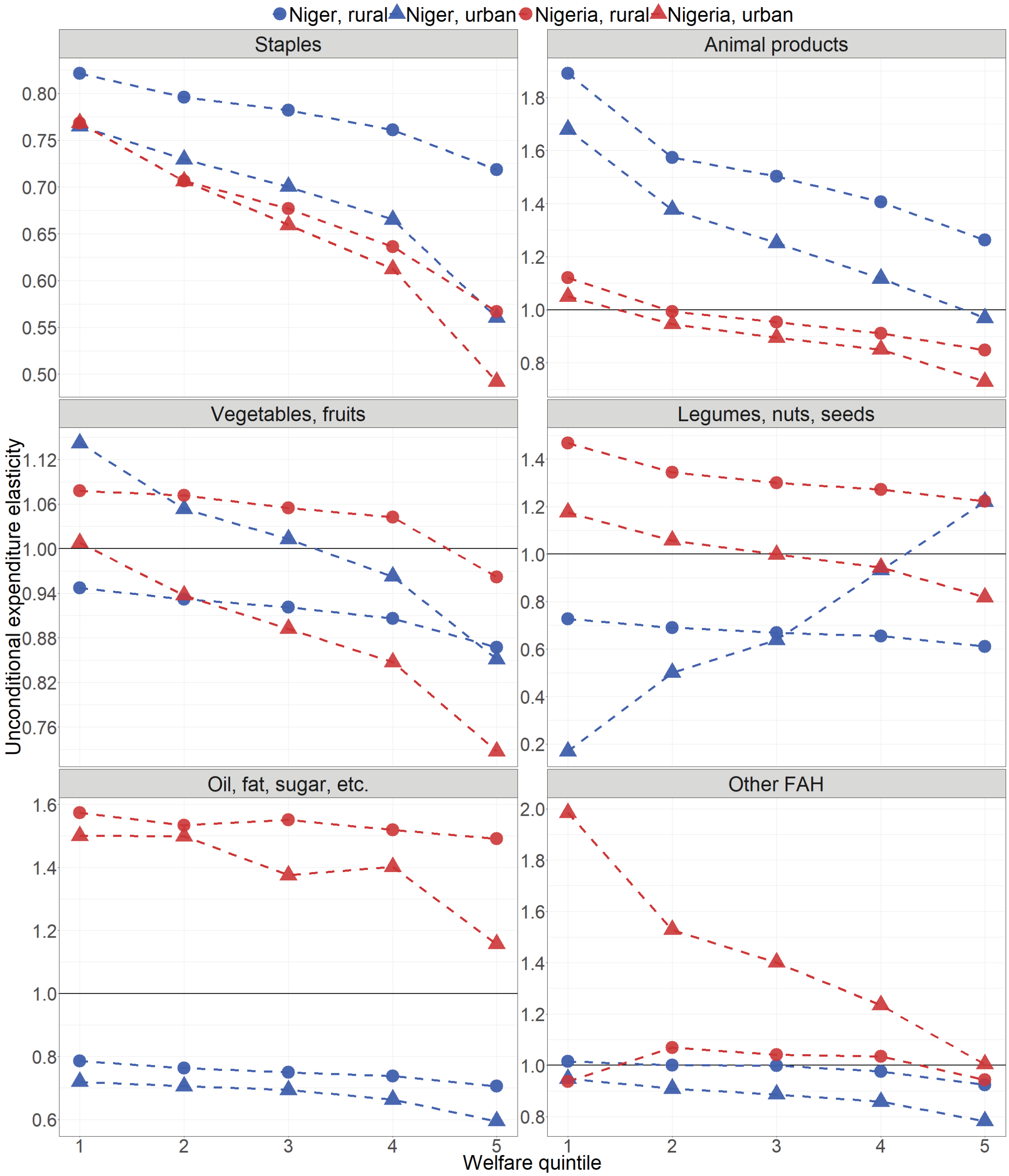


Figure 1.8. Second-stage own-price elasticities by welfare quintile

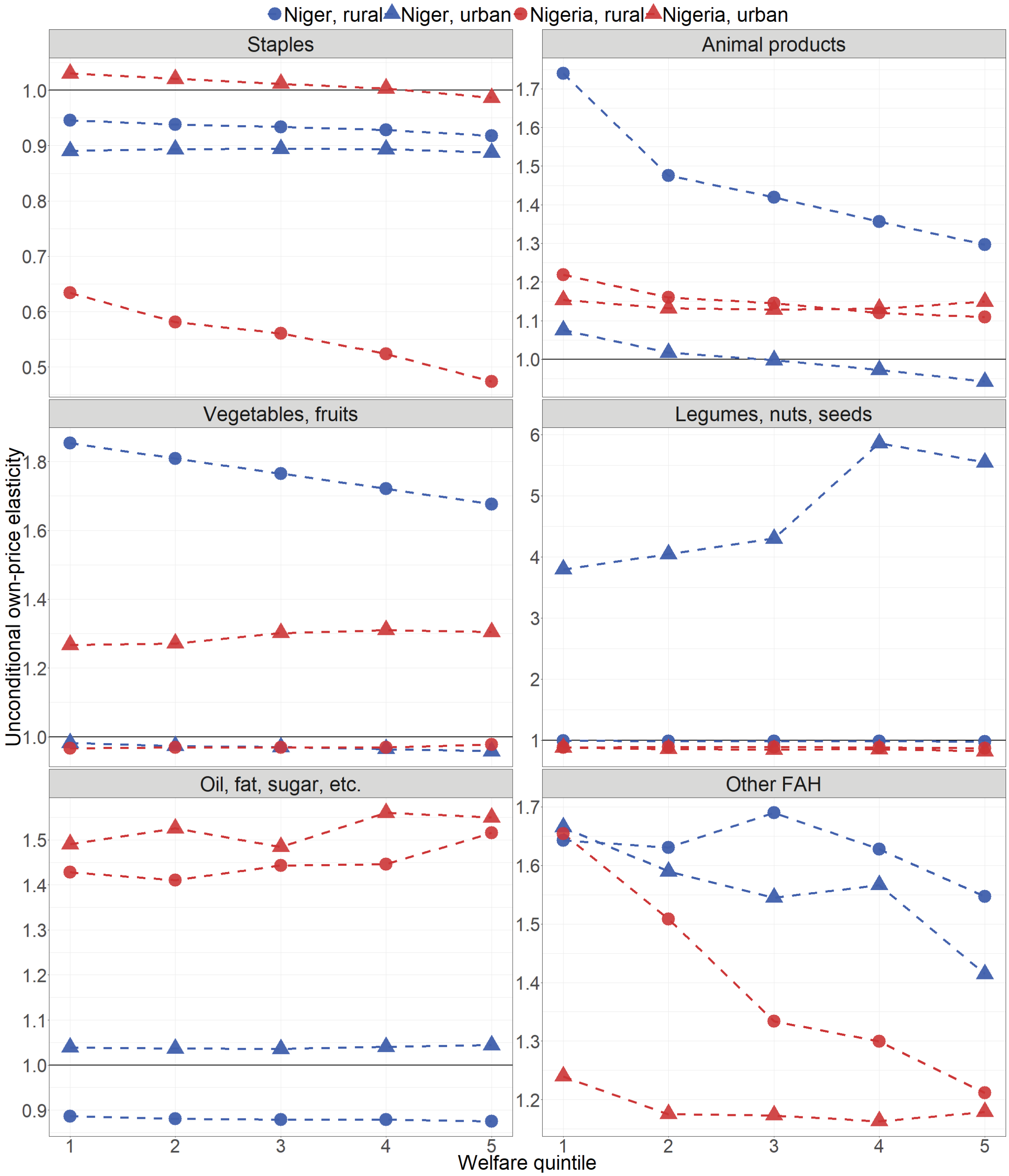


Figure 1.9. Third-stage expenditure elasticities by welfare quintile

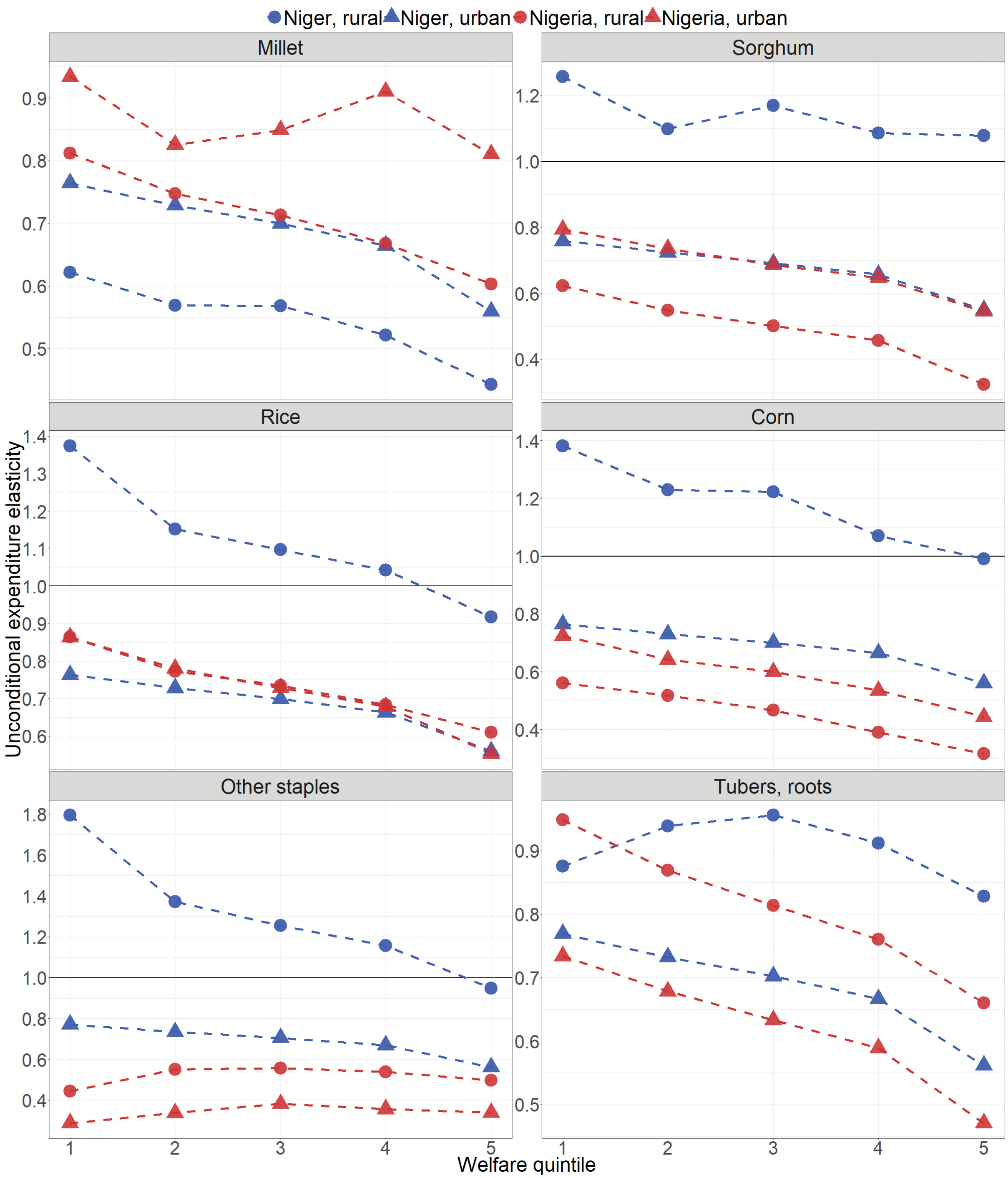


Figure 1.10. Third-stage own-price elasticities by welfare quintile

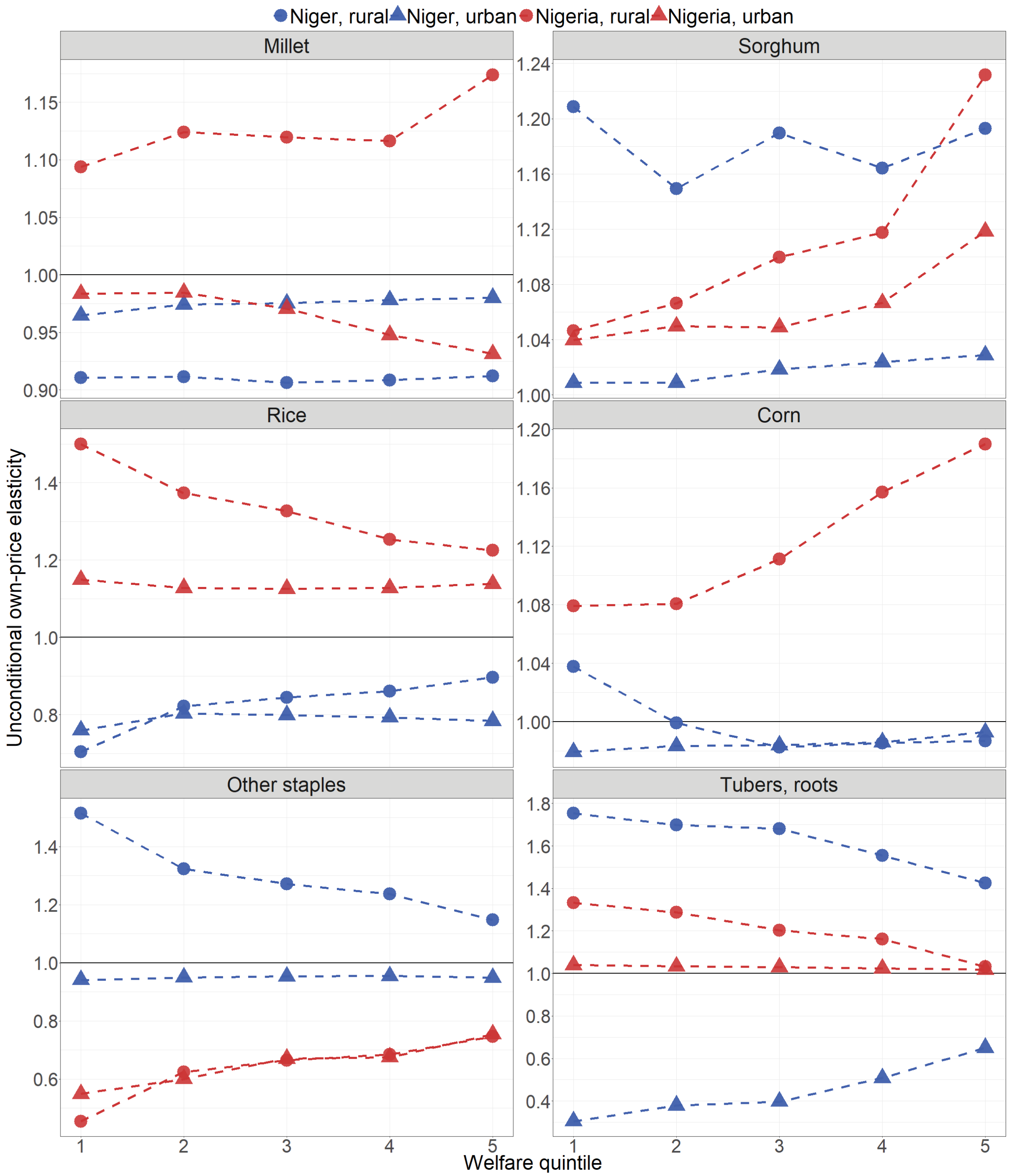
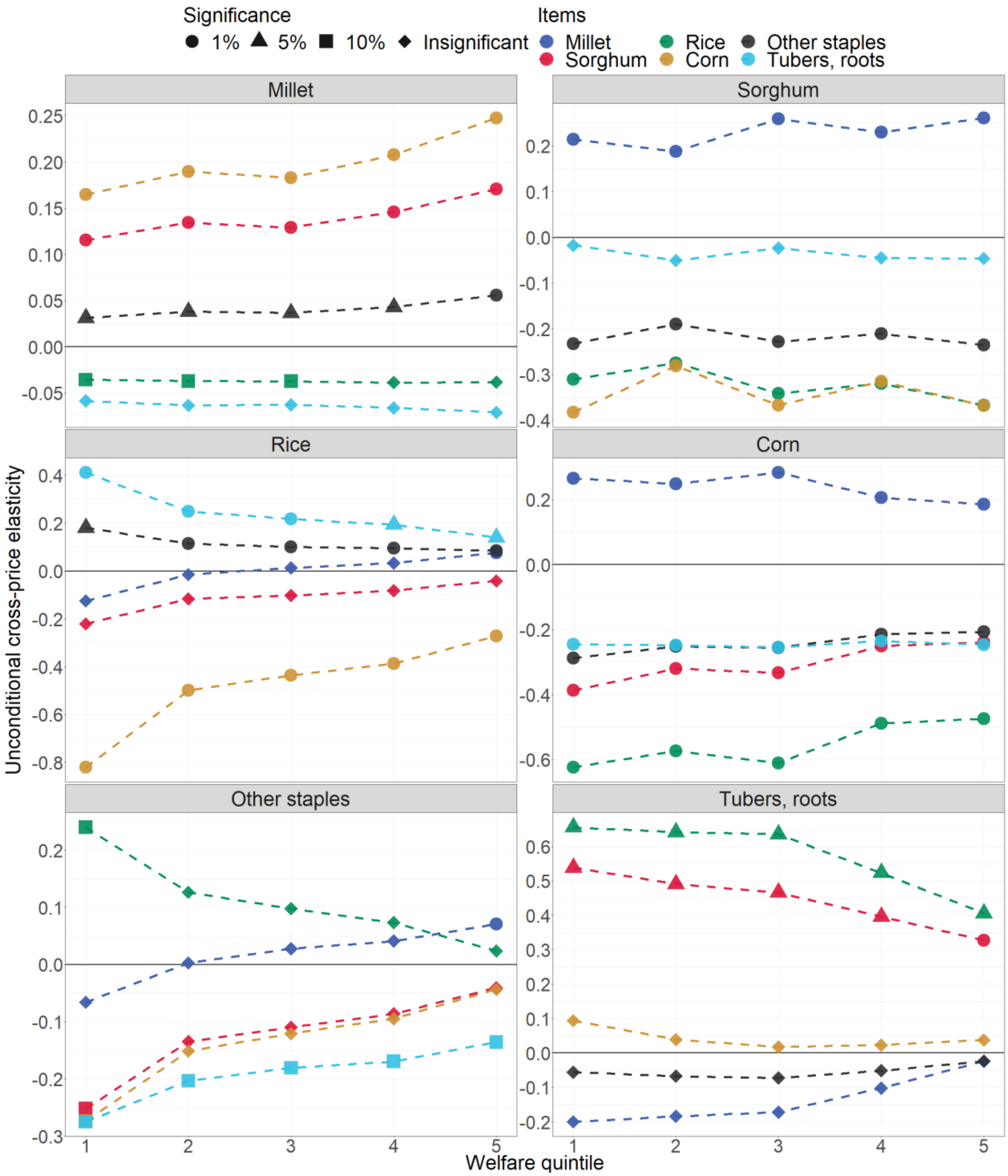


Figure 1.11. Third-stage cross-price elasticities by welfare quintile, rural Niger[13]



[13] The top-left panel shows the cross-price elasticities of the demand for millet with respect to the price of other subgroups. Other panels in Figure 1.11 and Figures 1.12-1.14 are interpreted similarly.

Figure 1.12. Third-stage cross-price elasticities by welfare quintile, urban Niger

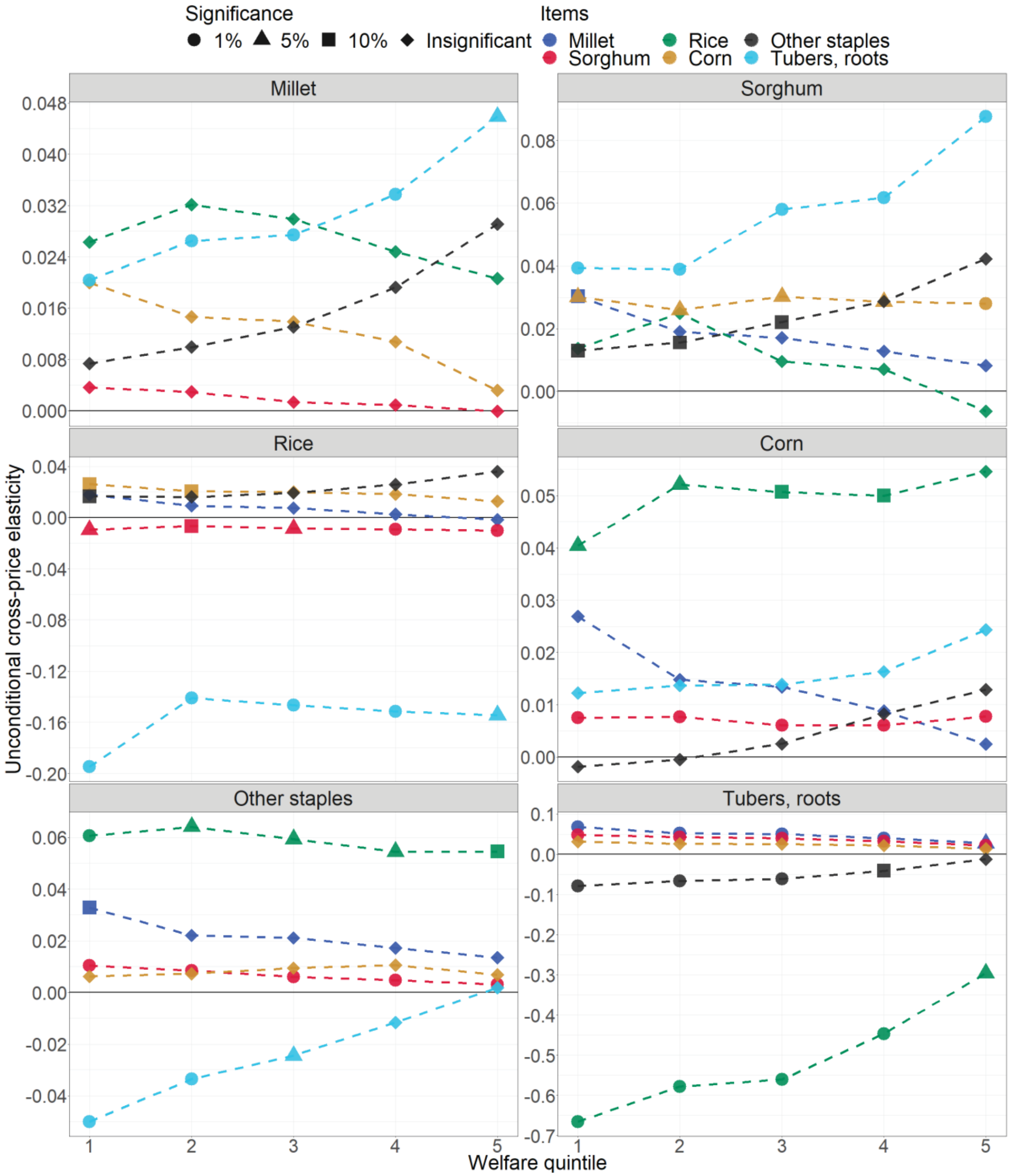


Figure 1.13. Third-stage cross-price elasticities by welfare quintile, rural Nigeria

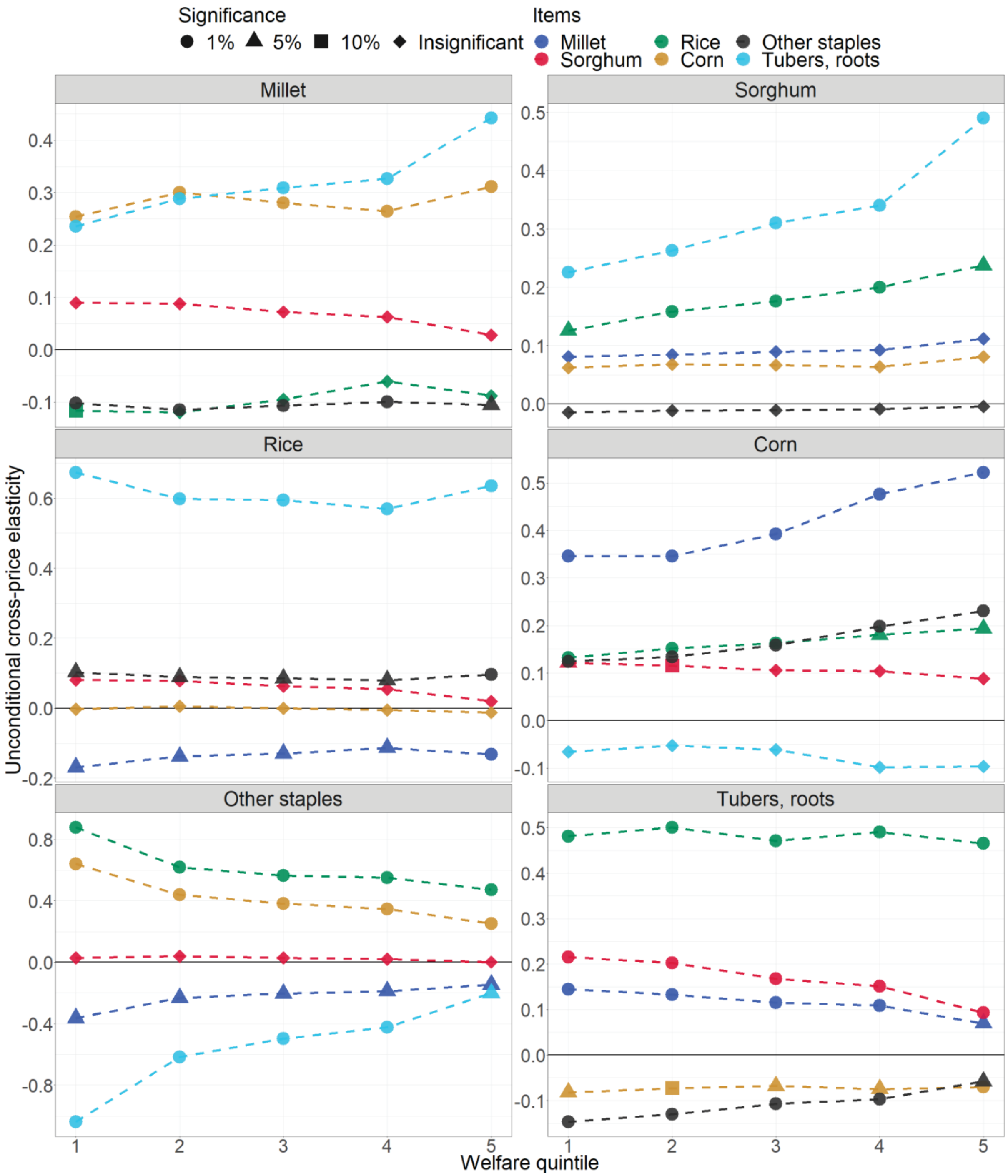


Figure 1.14. Third-stage cross-price elasticities by welfare quintile, urban Nigeria

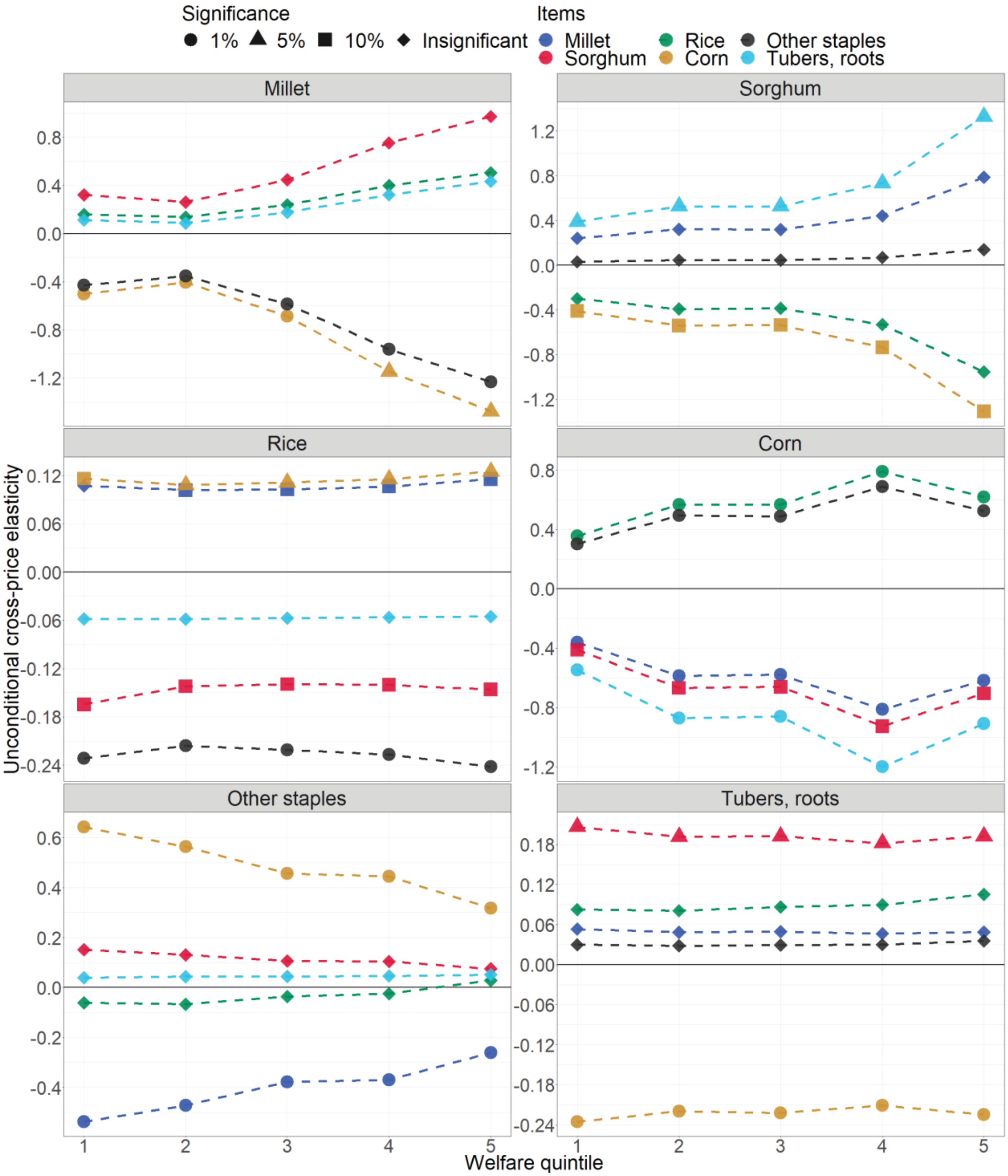


Figure 1.15. Marginal effects of HH characteristics on 2nd-stage budget shares, rural Niger (%)

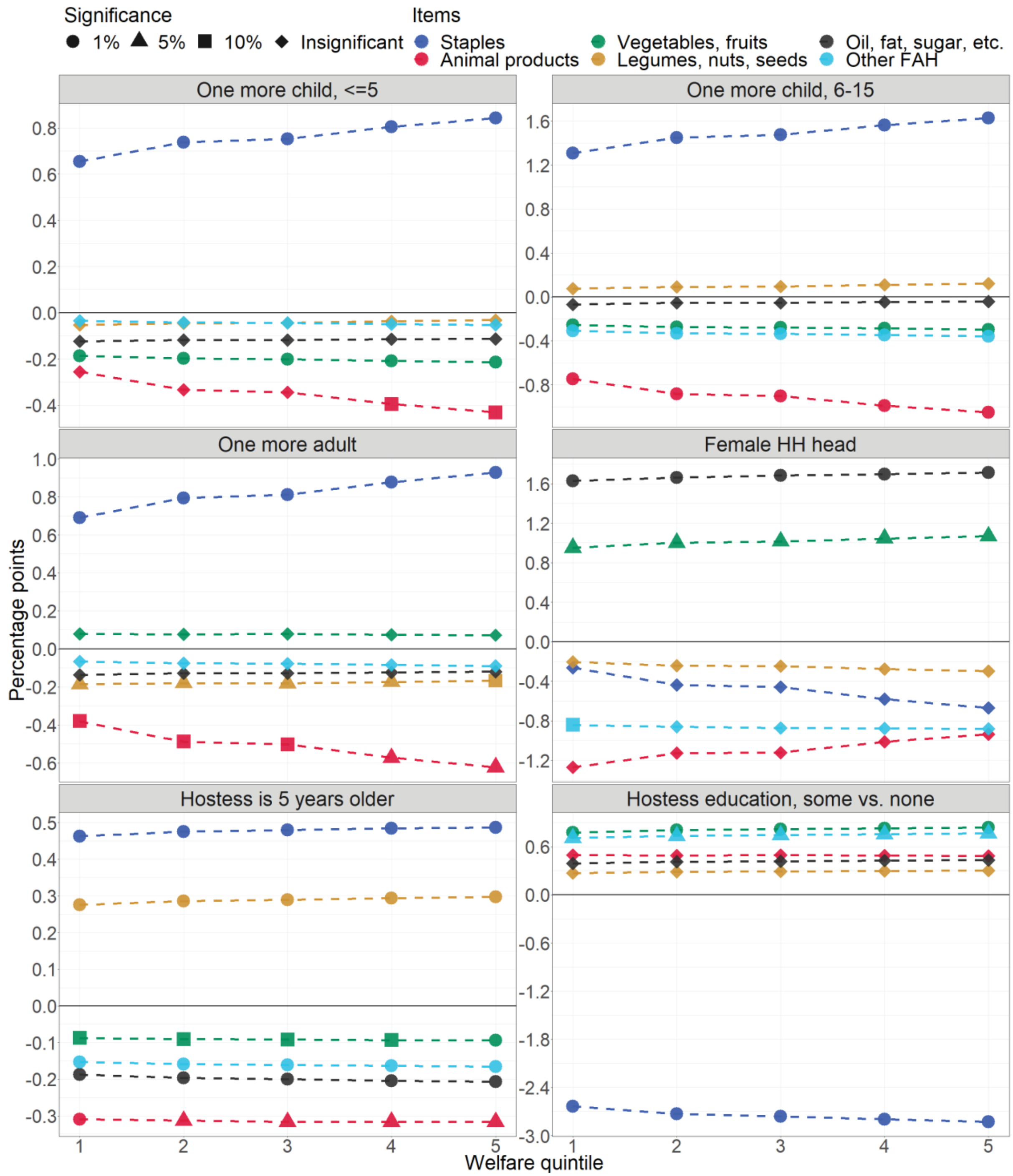


Figure 1.16. Marginal effects of HH characteristics on 2nd-stage budget shares, urban Niger (%)

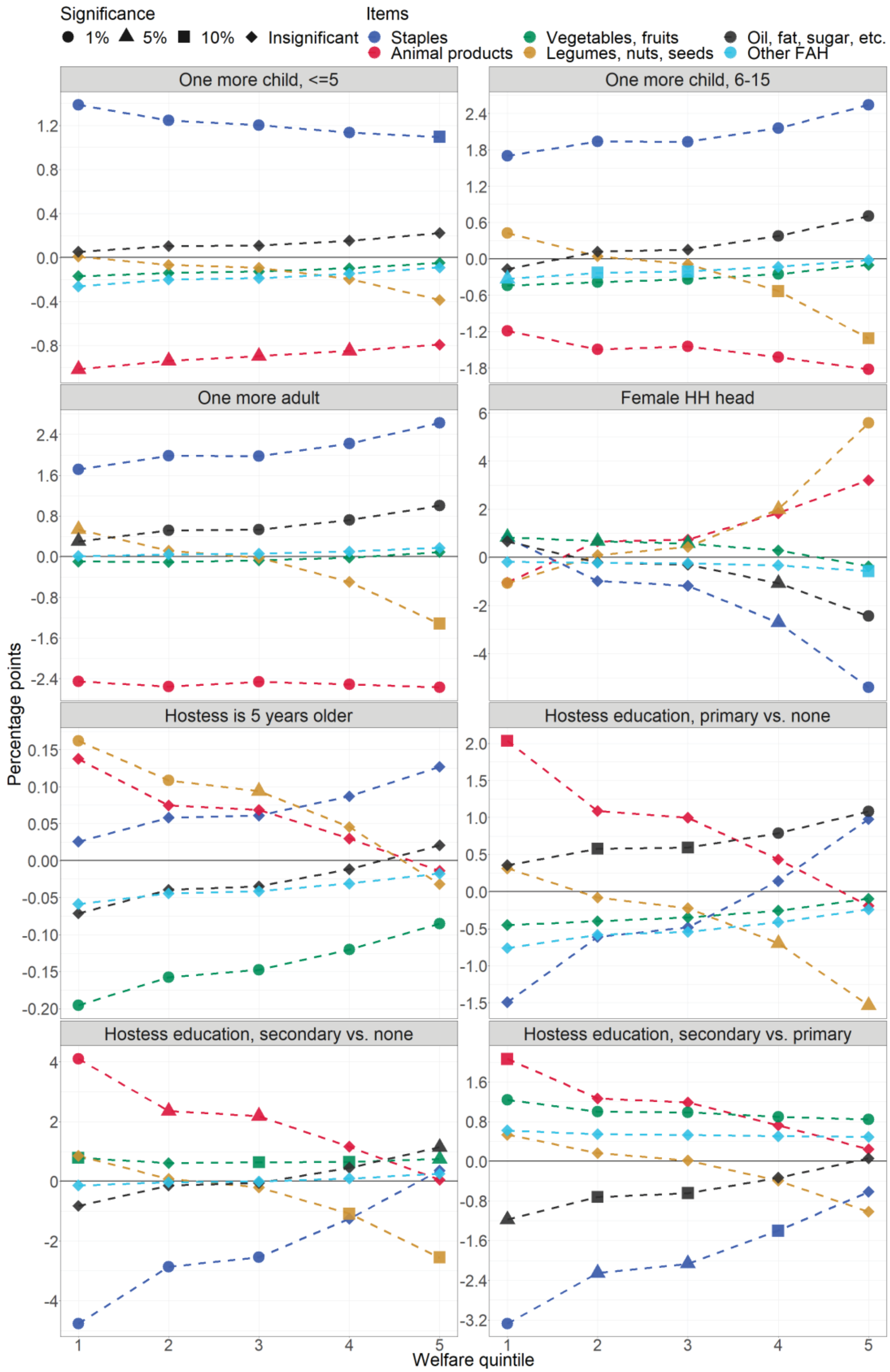


Figure 1.17. Marginal effects of HH characteristics on 2nd-stage budget shares, rural Nigeria (%)

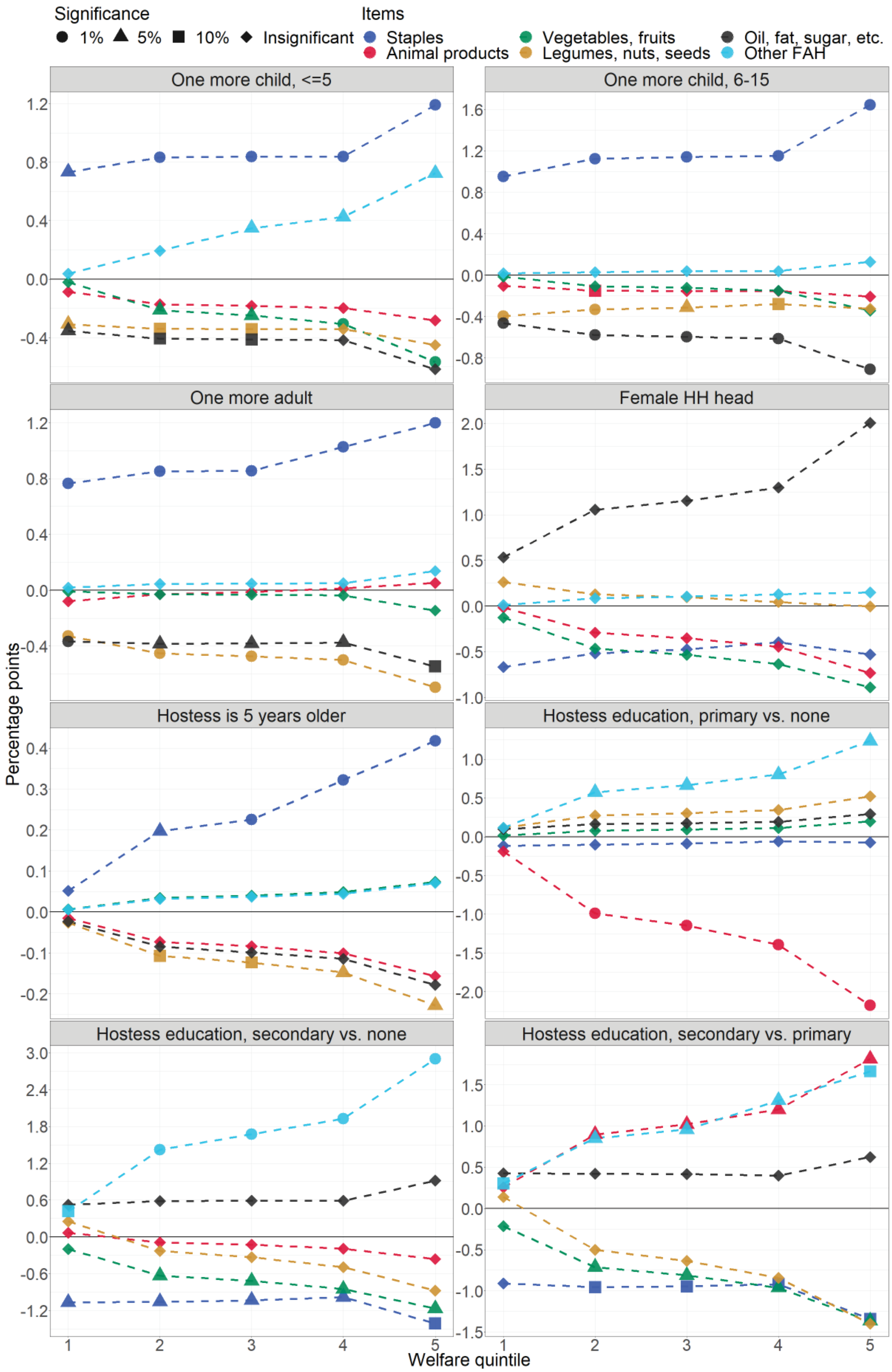


Figure 1.18. Marginal effects of HH characteristics on 2nd-stage budget shares, urban Nigeria (%)

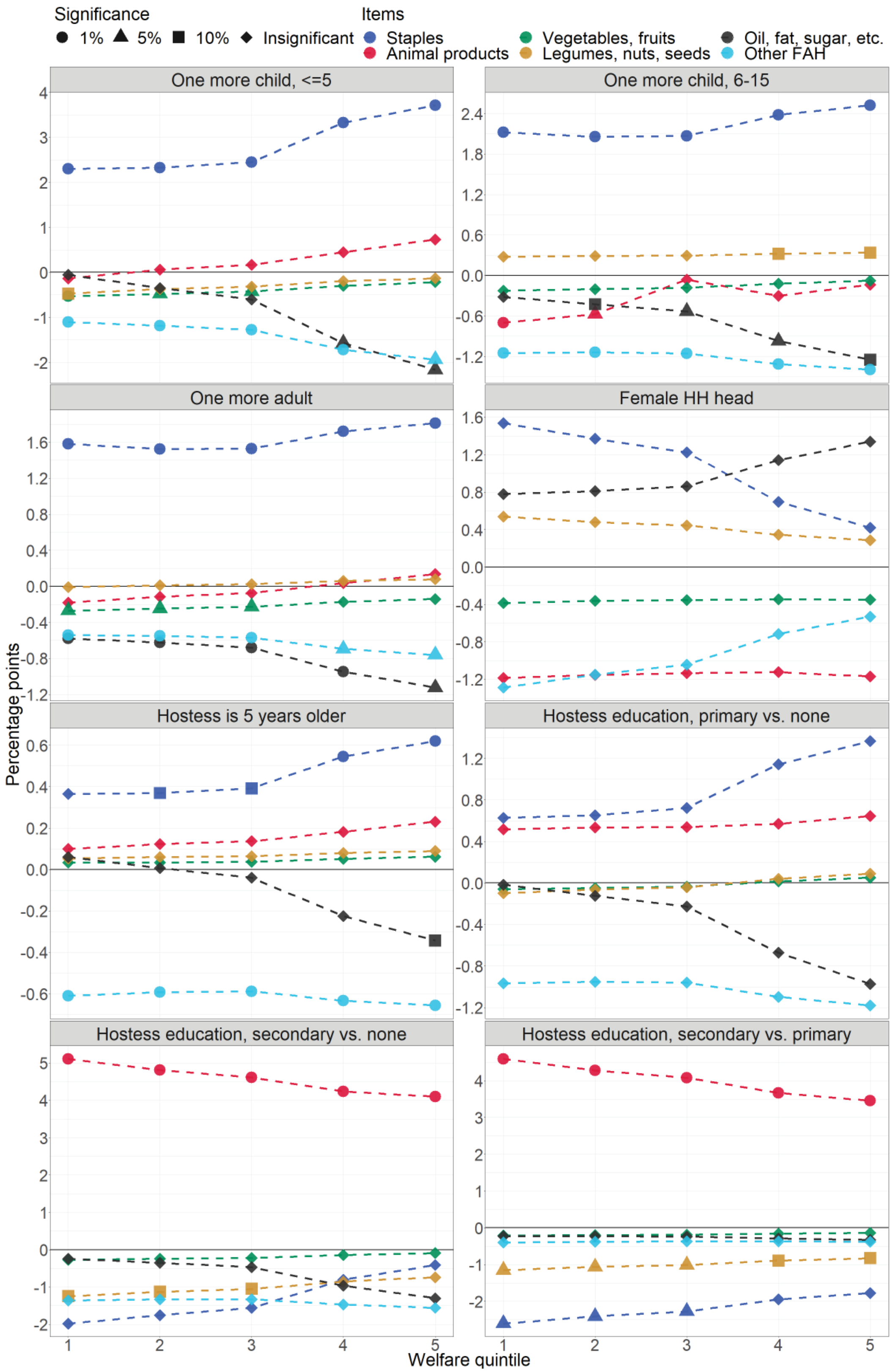


Figure 1.19. Marginal effects of HH characteristics on 3rd-stage budget shares, rural Niger (%)

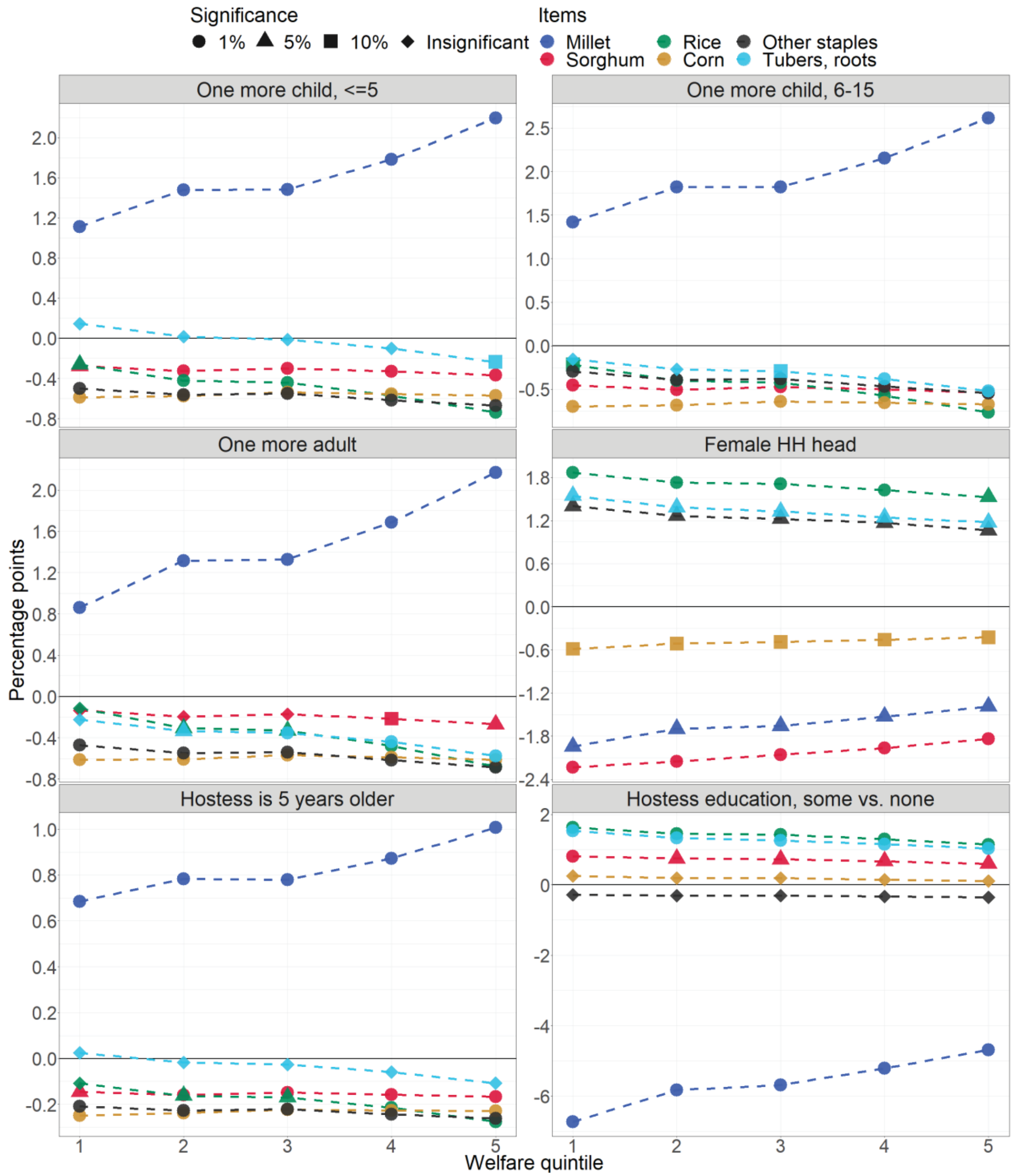


Figure 1.20. Marginal effects of HH characteristics on 3rd-stage budget shares, urban Niger (%)

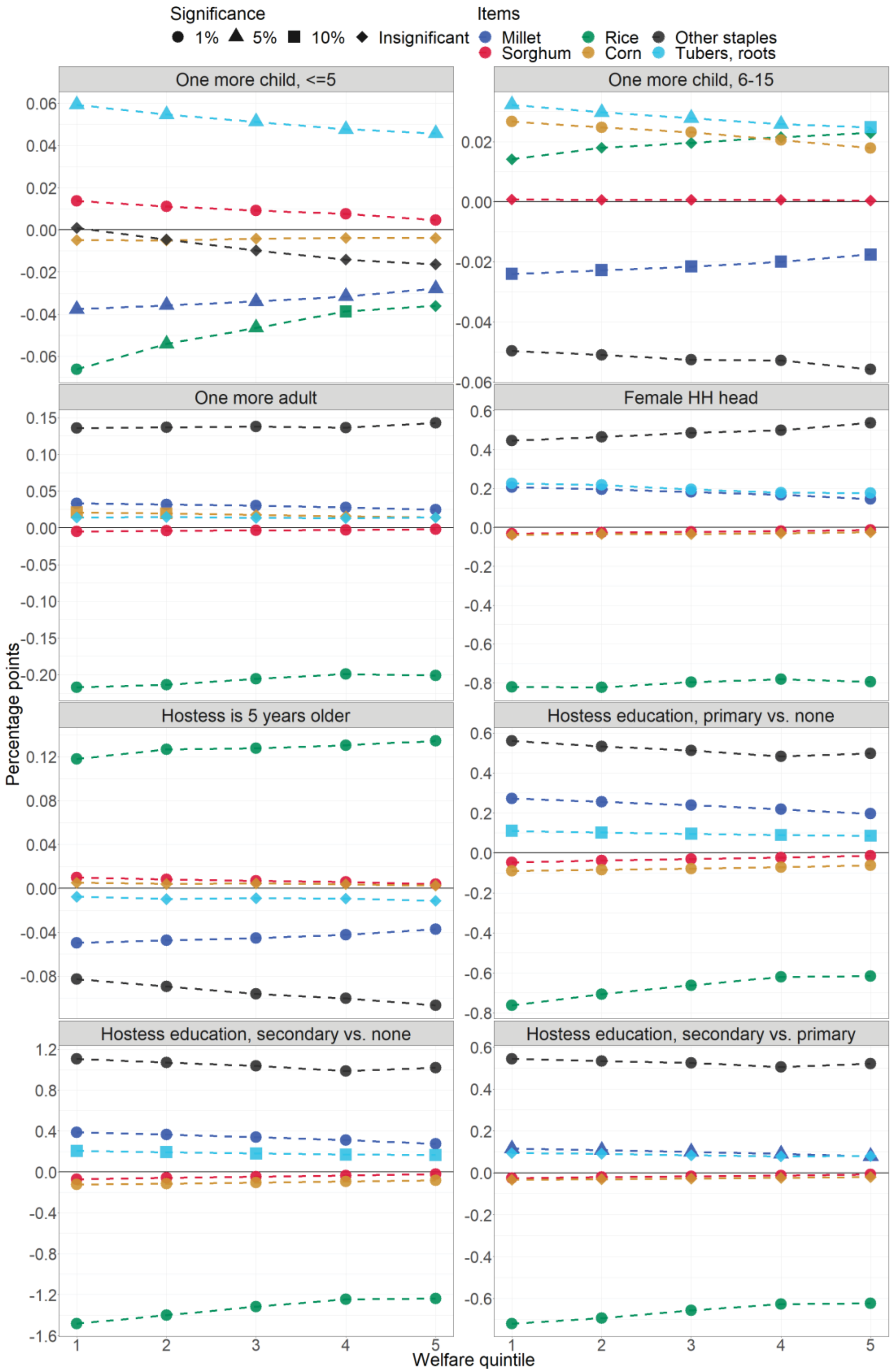


Figure 1.21. Marginal effects of HH characteristics on 3rd-stage budget shares, rural Nigeria (%)

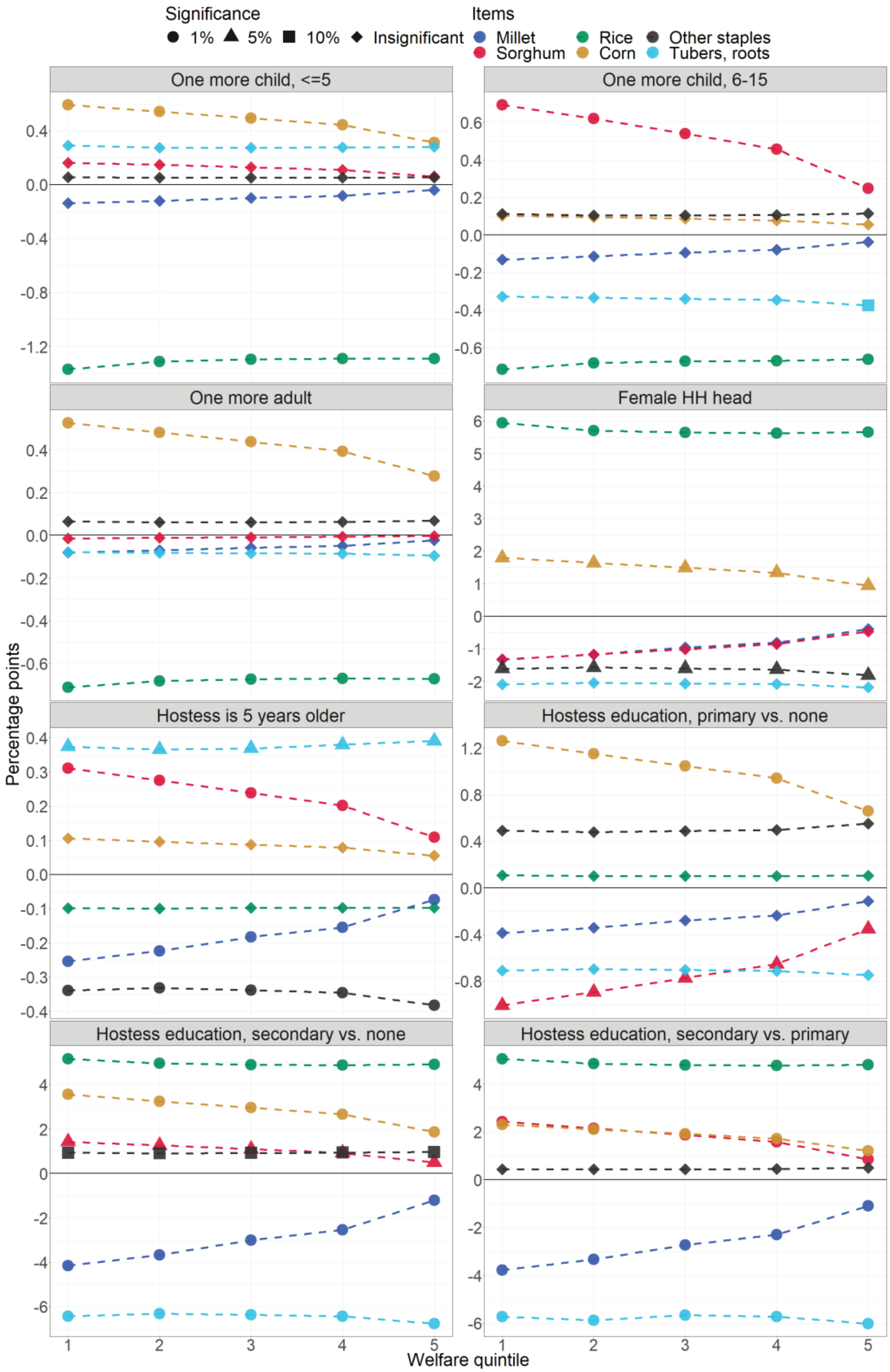
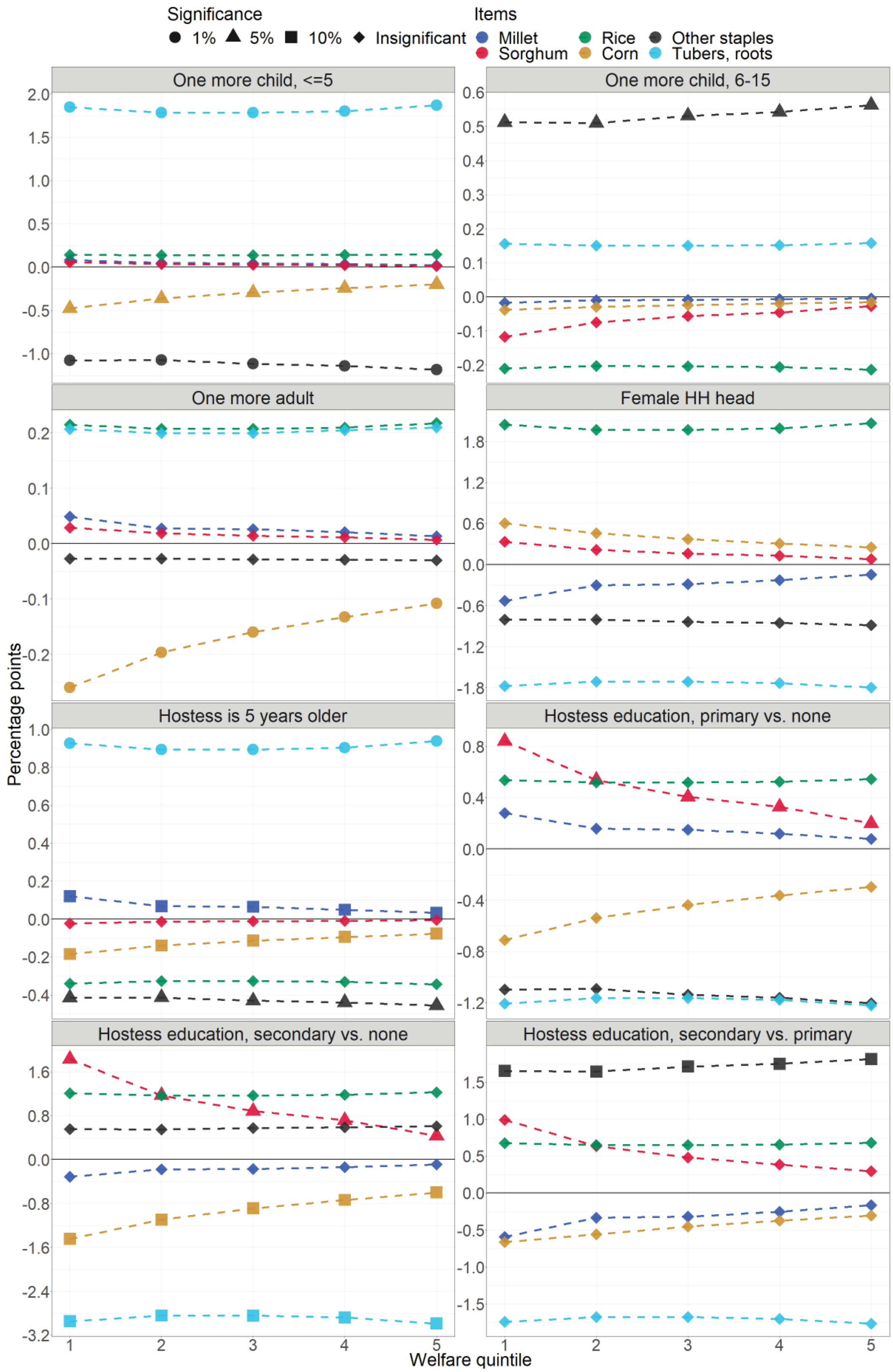


Figure 1.22. Marginal effects of HH characteristics on 3rd-stage budget shares, urban Nigeria (%)



Appendix tables

App. Table 1.1. Unconditional expenditure and uncompensated own-price elasticities by welfare quintile, first and second stages[14]

Food group	Quintile	Expenditure elasticity				Own-price elasticity			
		Niger		Nigeria		Niger		Nigeria	
		Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
FAH	1	0.911	0.906	1.032	1.014	-0.957	-0.868	-1.075	-1.293
FAH	2	0.895	0.885	0.994	0.950	-0.957	-0.863	-1.074	-1.308
FAH	3	0.884	0.869	0.970	0.908	-0.957	-0.857	-1.075	-1.332
FAH	4	0.871	0.848	0.942	0.863	-0.956	-0.848	-1.077	-1.341
FAH	5	0.835	0.780	0.878	0.741	-0.952	-0.811	-1.085	-1.408
1	1	0.821	0.765	0.768	0.769	-0.946	-0.890	-0.634	-1.030
1	2	0.796	0.730	0.706	0.706	-0.938	-0.893	-0.581	-1.021
1	3	0.782	0.700	0.677	0.659	-0.933	-0.894	-0.560	-1.012
1	4	0.761	0.665	0.636	0.612	-0.928	-0.893	-0.524	-1.003
1	5	0.718	0.560	0.567	0.492	-0.918	-0.887	-0.473	-0.987
2	1	1.891	1.680	1.121	1.049	-1.740	-1.075	-1.218	-1.154
2	2	1.574	1.377	0.993	0.946	-1.476	-1.017	-1.160	-1.132
2	3	1.503	1.251	0.954	0.894	-1.419	-0.998	-1.145	-1.129
2	4	1.406	1.118	0.910	0.850	-1.357	-0.972	-1.120	-1.131
2	5	1.263	0.969	0.848	0.729	-1.297	-0.942	-1.109	-1.149
3	1	0.947	1.142	1.078	1.008	-1.854	-0.982	-0.967	-1.267
3	2	0.932	1.054	1.072	0.937	-1.809	-0.973	-0.969	-1.271
3	3	0.922	1.013	1.055	0.892	-1.765	-0.970	-0.969	-1.302
3	4	0.906	0.963	1.042	0.847	-1.721	-0.965	-0.969	-1.310
3	5	0.867	0.852	0.962	0.727	-1.677	-0.958	-0.978	-1.305
4	1	0.728	0.169	1.467	1.176	-0.989	-3.794	-0.880	-0.878
4	2	0.691	0.500	1.344	1.057	-0.985	-4.044	-0.889	-0.854
4	3	0.669	0.639	1.300	0.998	-0.983	-4.298	-0.890	-0.843
4	4	0.656	0.932	1.271	0.943	-0.981	-5.858	-0.880	-0.850
4	5	0.611	1.221	1.223	0.818	-0.978	-5.545	-0.866	-0.816
5	1	0.786	0.719	1.573	1.500	-0.886	-1.039	-1.428	-1.491
5	2	0.763	0.706	1.533	1.498	-0.881	-1.037	-1.410	-1.526
5	3	0.749	0.693	1.551	1.375	-0.878	-1.035	-1.443	-1.484
5	4	0.738	0.664	1.519	1.401	-0.879	-1.041	-1.446	-1.560
5	5	0.705	0.595	1.491	1.157	-0.875	-1.044	-1.516	-1.550
6	1	1.016	0.947	0.936	1.984	-1.643	-1.665	-1.654	-1.240
6	2	1.000	0.909	1.070	1.528	-1.631	-1.590	-1.508	-1.175
6	3	0.999	0.887	1.041	1.401	-1.690	-1.545	-1.334	-1.173
6	4	0.976	0.858	1.035	1.234	-1.628	-1.566	-1.299	-1.163
6	5	0.925	0.782	0.943	1.004	-1.547	-1.415	-1.212	-1.179

[14] To make the appendix tables concise, the significance of the elasticities and marginal effects are represented by the background color: white (1%), green (5%), blue (10%), yellow (insignificant at 10%).

App. Table 1.2. Unconditional expenditure and uncompensated own-price elasticities by welfare quintile, the third stage

Staple subgroup	Quintile	Expenditure elasticity				Own-price elasticity			
		Niger		Nigeria		Niger		Nigeria	
		Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
1	1	0.622	0.764	0.812	0.934	-0.911	-0.965	-1.094	-0.984
1	2	0.569	0.729	0.748	0.826	-0.911	-0.974	-1.124	-0.985
1	3	0.568	0.700	0.713	0.849	-0.906	-0.975	-1.120	-0.971
1	4	0.522	0.665	0.668	0.910	-0.908	-0.978	-1.116	-0.948
1	5	0.443	0.560	0.603	0.810	-0.912	-0.980	-1.174	-0.931
2	1	1.257	0.759	0.623	0.794	-1.209	-1.009	-1.046	-1.040
2	2	1.098	0.724	0.549	0.735	-1.149	-1.009	-1.066	-1.050
2	3	1.170	0.691	0.501	0.686	-1.190	-1.019	-1.100	-1.049
2	4	1.085	0.657	0.457	0.647	-1.164	-1.024	-1.118	-1.067
2	5	1.078	0.550	0.323	0.545	-1.193	-1.029	-1.232	-1.118
3	1	1.374	0.764	0.864	0.864	-0.704	-0.760	-1.499	-1.149
3	2	1.152	0.728	0.773	0.780	-0.822	-0.804	-1.374	-1.128
3	3	1.098	0.699	0.735	0.728	-0.845	-0.800	-1.327	-1.126
3	4	1.042	0.663	0.684	0.679	-0.861	-0.793	-1.253	-1.128
3	5	0.918	0.559	0.610	0.552	-0.896	-0.785	-1.226	-1.139
4	1	1.381	0.765	0.561	0.725	-1.038	-0.979	-1.079	-0.291
4	2	1.230	0.730	0.518	0.642	-0.999	-0.983	-1.081	0.150
4	3	1.223	0.700	0.467	0.600	-0.983	-0.984	-1.111	0.137
4	4	1.070	0.665	0.391	0.535	-0.985	-0.986	-1.157	0.594
4	5	0.990	0.561	0.316	0.444	-0.987	-0.993	-1.190	0.214
5	1	1.794	0.771	0.446	0.288	-1.514	-0.941	-0.455	-0.548
5	2	1.372	0.735	0.552	0.338	-1.323	-0.950	-0.624	-0.600
5	3	1.254	0.705	0.557	0.383	-1.272	-0.953	-0.664	-0.670
5	4	1.156	0.669	0.538	0.357	-1.237	-0.954	-0.684	-0.675
5	5	0.948	0.563	0.498	0.340	-1.147	-0.949	-0.747	-0.754
6	1	0.875	0.769	0.949	0.734	-1.753	-0.305	-1.333	-1.039
6	2	0.939	0.732	0.869	0.679	-1.699	-0.379	-1.287	-1.033
6	3	0.956	0.702	0.814	0.633	-1.680	-0.397	-1.202	-1.030
6	4	0.912	0.666	0.761	0.589	-1.555	-0.508	-1.162	-1.024
6	5	0.828	0.561	0.660	0.471	-1.424	-0.651	-1.032	-1.018

App. Table 1.3. Cross-price elasticities by welfare quintile, the third stage

Group, demand	Group, price	Quintile	Niger		Nigeria	
			Rural	Urban	Rural	Urban
1	2	1	0.115	0.004	0.090	0.320
1	2	2	0.135	0.003	0.088	0.261
1	2	3	0.129	0.001	0.072	0.448
1	2	4	0.146	0.001	0.063	0.753
1	2	5	0.170	0.000	0.028	0.971
1	3	1	-0.036	0.026	-0.116	0.159
1	3	2	-0.037	0.032	-0.120	0.136
1	3	3	-0.037	0.030	-0.094	0.238
1	3	4	-0.039	0.025	-0.060	0.399
1	3	5	-0.038	0.021	-0.088	0.505
1	4	1	0.165	0.020	0.254	-0.504
1	4	2	0.190	0.015	0.300	-0.406
1	4	3	0.183	0.014	0.280	-0.685
1	4	4	0.208	0.011	0.265	-1.145
1	4	5	0.248	0.003	0.311	-1.476
1	5	1	0.031	0.007	-0.102	-0.430
1	5	2	0.038	0.010	-0.114	-0.354
1	5	3	0.037	0.013	-0.106	-0.584
1	5	4	0.043	0.019	-0.099	-0.960
1	5	5	0.056	0.029	-0.105	-1.230
1	6	1	-0.059	0.020	0.235	0.114
1	6	2	-0.063	0.026	0.288	0.087
1	6	3	-0.063	0.027	0.309	0.176
1	6	4	-0.066	0.034	0.326	0.321
1	6	5	-0.071	0.046	0.442	0.432
2	1	1	0.215	0.030	0.081	0.241
2	1	2	0.188	0.019	0.084	0.321
2	1	3	0.260	0.017	0.089	0.320
2	1	4	0.230	0.013	0.093	0.441
2	1	5	0.261	0.008	0.111	0.786
2	3	1	-0.310	0.014	0.125	-0.300
2	3	2	-0.275	0.025	0.158	-0.391
2	3	3	-0.342	0.010	0.176	-0.384
2	3	4	-0.320	0.007	0.200	-0.529
2	3	5	-0.367	-0.006	0.237	-0.953
2	4	1	-0.383	0.030	0.062	-0.410
2	4	2	-0.281	0.026	0.068	-0.539
2	4	3	-0.367	0.030	0.067	-0.535
2	4	4	-0.315	0.029	0.064	-0.735
2	4	5	-0.368	0.028	0.081	-1.307
2	5	1	-0.233	0.013	-0.015	0.030
2	5	2	-0.190	0.016	-0.012	0.045
2	5	3	-0.228	0.022	-0.011	0.044
2	5	4	-0.211	0.029	-0.009	0.068
2	5	5	-0.236	0.042	-0.004	0.139
2	6	1	-0.017	0.039	0.226	0.389
2	6	2	-0.050	0.039	0.263	0.526

2	6	3	-0.023	0.058	0.310	0.525
2	6	4	-0.045	0.062	0.340	0.734
2	6	5	-0.046	0.088	0.490	1.329
3	1	1	-0.125	0.018	-0.168	0.107
3	1	2	-0.014	0.009	-0.138	0.102
3	1	3	0.014	0.007	-0.129	0.103
3	1	4	0.034	0.003	-0.113	0.106
3	1	5	0.076	-0.002	-0.132	0.115
3	2	1	-0.221	-0.010	0.081	-0.165
3	2	2	-0.116	-0.007	0.079	-0.142
3	2	3	-0.102	-0.009	0.063	-0.139
3	2	4	-0.082	-0.009	0.054	-0.140
3	2	5	-0.042	-0.010	0.020	-0.146
3	4	1	-0.820	0.026	-0.003	0.116
3	4	2	-0.498	0.021	0.006	0.108
3	4	3	-0.436	0.020	0.000	0.111
3	4	4	-0.386	0.018	-0.005	0.116
3	4	5	-0.272	0.013	-0.012	0.125
3	5	1	0.181	0.017	0.103	-0.231
3	5	2	0.115	0.016	0.089	-0.216
3	5	3	0.101	0.019	0.086	-0.221
3	5	4	0.096	0.026	0.080	-0.227
3	5	5	0.085	0.036	0.096	-0.242
3	6	1	0.411	-0.195	0.673	-0.058
3	6	2	0.249	-0.141	0.599	-0.058
3	6	3	0.217	-0.146	0.595	-0.057
3	6	4	0.194	-0.151	0.570	-0.056
3	6	5	0.141	-0.154	0.635	-0.055
4	1	1	0.265	0.027	0.346	-0.362
4	1	2	0.249	0.015	0.346	-0.587
4	1	3	0.283	0.013	0.392	-0.580
4	1	4	0.206	0.009	0.476	-0.812
4	1	5	0.185	0.002	0.521	-0.618
4	2	1	-0.387	0.008	0.122	-0.412
4	2	2	-0.320	0.008	0.115	-0.668
4	2	3	-0.332	0.006	0.106	-0.661
4	2	4	-0.250	0.006	0.104	-0.927
4	2	5	-0.240	0.008	0.088	-0.705
4	3	1	-0.623	0.040	0.132	0.354
4	3	2	-0.574	0.052	0.152	0.568
4	3	3	-0.611	0.051	0.163	0.566
4	3	4	-0.488	0.050	0.180	0.789
4	3	5	-0.474	0.055	0.194	0.616
4	5	1	-0.288	-0.002	0.124	0.301
4	5	2	-0.252	0.000	0.134	0.494
4	5	3	-0.256	0.003	0.159	0.488
4	5	4	-0.215	0.008	0.198	0.689
4	5	5	-0.207	0.013	0.231	0.524
4	6	1	-0.245	0.012	-0.066	-0.547
4	6	2	-0.249	0.014	-0.053	-0.873
4	6	3	-0.254	0.014	-0.062	-0.860

4	6	4	-0.236	0.016	-0.098	-1.198
4	6	5	-0.247	0.024	-0.097	-0.909
5	1	1	-0.066	0.033	-0.364	-0.538
5	1	2	0.003	0.022	-0.232	-0.473
5	1	3	0.028	0.021	-0.204	-0.379
5	1	4	0.041	0.017	-0.189	-0.371
5	1	5	0.071	0.013	-0.146	-0.261
5	2	1	-0.252	0.010	0.027	0.151
5	2	2	-0.134	0.008	0.038	0.131
5	2	3	-0.110	0.006	0.028	0.106
5	2	4	-0.086	0.005	0.020	0.104
5	2	5	-0.040	0.003	0.001	0.075
5	3	1	0.240	0.061	0.879	-0.061
5	3	2	0.127	0.064	0.618	-0.067
5	3	3	0.098	0.059	0.565	-0.037
5	3	4	0.073	0.054	0.551	-0.025
5	3	5	0.024	0.054	0.473	0.029
5	4	1	-0.273	0.006	0.641	0.643
5	4	2	-0.151	0.007	0.439	0.564
5	4	3	-0.121	0.009	0.382	0.456
5	4	4	-0.094	0.011	0.347	0.445
5	4	5	-0.043	0.007	0.253	0.318
5	6	1	-0.275	-0.050	-1.040	0.039
5	6	2	-0.203	-0.034	-0.617	0.044
5	6	3	-0.181	-0.025	-0.497	0.044
5	6	4	-0.169	-0.012	-0.423	0.046
5	6	5	-0.136	0.002	-0.201	0.051
6	1	1	-0.200	0.068	0.145	0.053
6	1	2	-0.184	0.052	0.132	0.049
6	1	3	-0.172	0.050	0.115	0.049
6	1	4	-0.102	0.039	0.108	0.047
6	1	5	-0.024	0.027	0.070	0.049
6	2	1	0.539	0.048	0.215	0.206
6	2	2	0.491	0.043	0.202	0.192
6	2	3	0.466	0.040	0.168	0.192
6	2	4	0.396	0.032	0.151	0.182
6	2	5	0.327	0.022	0.093	0.192
6	3	1	0.656	-0.665	0.481	0.082
6	3	2	0.642	-0.578	0.501	0.081
6	3	3	0.636	-0.560	0.471	0.086
6	3	4	0.523	-0.447	0.490	0.090
6	3	5	0.406	-0.296	0.466	0.105
6	4	1	0.094	0.031	-0.082	-0.235
6	4	2	0.039	0.026	-0.073	-0.220
6	4	3	0.017	0.025	-0.068	-0.222
6	4	4	0.023	0.022	-0.075	-0.211
6	4	5	0.038	0.014	-0.071	-0.225
6	5	1	-0.056	-0.079	-0.147	0.030
6	5	2	-0.067	-0.067	-0.130	0.028
6	5	3	-0.072	-0.062	-0.107	0.030
6	5	4	-0.051	-0.041	-0.097	0.030

6	5	5	-0.023	-0.013	-0.058	0.035

Note: the elasticity in this table is the elasticity of the demand for "Group, demand" with respect to the price of "Group, price".

App. Table 1.4. Marginal effects of household demographic characteristics by welfare quintile

Food group or staple subgroup	Quintile	Second stage				Third stage			
		Niger		Nigeria		Niger		Nigeria	
		Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
One more child, <=5									
1	1	0.655	1.389	0.733	2.298	1.111	-0.038	-0.138	0.084
1	2	0.738	1.247	0.833	2.325	1.478	-0.036	-0.121	0.047
1	3	0.752	1.202	0.839	2.453	1.481	-0.034	-0.098	0.045
1	4	0.804	1.136	0.839	3.330	1.782	-0.032	-0.083	0.036
1	5	0.843	1.097	1.193	3.718	2.195	-0.028	-0.040	0.023
2	1	-0.256	-1.017	-0.087	-0.140	-0.275	0.014	0.161	0.055
2	2	-0.334	-0.940	-0.172	0.058	-0.324	0.011	0.148	0.035
2	3	-0.345	-0.898	-0.184	0.168	-0.298	0.009	0.129	0.027
2	4	-0.396	-0.849	-0.198	0.451	-0.330	0.008	0.110	0.022
2	5	-0.432	-0.793	-0.282	0.732	-0.369	0.005	0.060	0.013
3	1	-0.187	-0.170	-0.021	-0.530	-0.262	-0.066	-1.369	0.145
3	2	-0.198	-0.142	-0.212	-0.478	-0.421	-0.054	-1.311	0.139
3	3	-0.201	-0.129	-0.249	-0.426	-0.442	-0.046	-1.296	0.139
3	4	-0.208	-0.098	-0.307	-0.297	-0.569	-0.039	-1.290	0.141
3	5	-0.214	-0.049	-0.567	-0.216	-0.739	-0.036	-1.291	0.146
4	1	-0.053	0.008	-0.308	-0.476	-0.585	-0.005	0.592	-0.477
4	2	-0.045	-0.067	-0.339	-0.371	-0.574	-0.005	0.542	-0.361
4	3	-0.044	-0.094	-0.342	-0.321	-0.538	-0.004	0.492	-0.293
4	4	-0.037	-0.196	-0.342	-0.194	-0.553	-0.004	0.443	-0.243
4	5	-0.032	-0.387	-0.451	-0.134	-0.572	-0.004	0.311	-0.198
5	1	-0.124	0.052	-0.353	-0.052	-0.501	0.001	0.055	-1.077
5	2	-0.118	0.103	-0.408	-0.350	-0.562	-0.005	0.050	-1.073
5	3	-0.119	0.107	-0.413	-0.602	-0.551	-0.010	0.050	-1.117
5	4	-0.115	0.152	-0.417	-1.572	-0.618	-0.014	0.050	-1.140
5	5	-0.113	0.223	-0.617	-2.162	-0.672	-0.016	0.054	-1.183
6	1	-0.036	-0.262	0.037	-1.101	0.144	0.059	0.289	1.849
6	2	-0.043	-0.201	0.193	-1.184	0.017	0.055	0.273	1.783
6	3	-0.043	-0.188	0.349	-1.273	-0.012	0.051	0.274	1.783
6	4	-0.049	-0.146	0.425	-1.718	-0.102	0.048	0.276	1.804
6	5	-0.052	-0.091	0.723	-1.938	-0.237	0.046	0.279	1.872
One more child, 6-15									
1	1	1.309	1.701	0.954	2.120	1.420	-0.024	-0.132	-0.017
1	2	1.449	1.941	1.126	2.055	1.822	-0.023	-0.114	-0.010
1	3	1.476	1.930	1.142	2.067	1.822	-0.022	-0.093	-0.009
1	4	1.563	2.159	1.154	2.382	2.156	-0.020	-0.079	-0.007
1	5	1.628	2.545	1.647	2.525	2.615	-0.018	-0.037	-0.005

	2	1	-0.747	-1.191	-0.104	-0.702	-0.454	0.001	0.694	-0.117
	2	2	-0.883	-1.492	-0.151	-0.573	-0.506	0.001	0.621	-0.075
	2	3	-0.904	-1.449	-0.154	-0.060	-0.471	0.001	0.539	-0.057
	2	4	-0.990	-1.625	-0.153	-0.303	-0.501	0.001	0.458	-0.046
	2	5	-1.052	-1.823	-0.205	-0.138	-0.539	0.000	0.249	-0.028
	3	1	-0.257	-0.437	-0.012	-0.227	-0.216	0.014	-0.715	-0.211
	3	2	-0.273	-0.382	-0.105	-0.204	-0.402	0.018	-0.681	-0.203
	3	3	-0.277	-0.338	-0.121	-0.181	-0.426	0.020	-0.673	-0.203
	3	4	-0.287	-0.250	-0.149	-0.119	-0.572	0.022	-0.670	-0.206
	3	5	-0.296	-0.100	-0.342	-0.079	-0.767	0.023	-0.662	-0.214
	4	1	0.074	0.426	-0.396	0.276	-0.698	0.027	0.102	-0.038
	4	2	0.092	0.045	-0.330	0.290	-0.681	0.025	0.095	-0.029
	4	3	0.095	-0.087	-0.312	0.293	-0.639	0.023	0.086	-0.024
	4	4	0.109	-0.534	-0.278	0.321	-0.655	0.021	0.078	-0.020
	4	5	0.120	-1.312	-0.322	0.336	-0.675	0.018	0.055	-0.016
	5	1	-0.069	-0.169	-0.463	-0.317	-0.296	-0.050	0.112	0.512
	5	2	-0.055	0.120	-0.577	-0.427	-0.388	-0.051	0.105	0.510
	5	3	-0.055	0.153	-0.594	-0.530	-0.383	-0.053	0.106	0.531
	5	4	-0.046	0.380	-0.613	-0.970	-0.466	-0.053	0.106	0.542
	5	5	-0.042	0.707	-0.908	-1.249	-0.544	-0.056	0.116	0.563
	6	1	-0.309	-0.329	0.020	-1.151	-0.155	0.032	-0.328	0.156
	6	2	-0.330	-0.231	0.029	-1.141	-0.271	0.030	-0.335	0.150
	6	3	-0.335	-0.210	0.038	-1.154	-0.292	0.028	-0.340	0.150
	6	4	-0.348	-0.130	0.038	-1.311	-0.380	0.026	-0.346	0.152
	6	5	-0.358	-0.018	0.131	-1.395	-0.521	0.025	-0.375	0.158
One more adult										
	1	1	0.689	1.716	0.766	1.581	0.860	0.033	-0.081	0.049
	1	2	0.793	1.980	0.852	1.525	1.317	0.032	-0.071	0.028
	1	3	0.810	1.972	0.854	1.526	1.328	0.030	-0.058	0.026
	1	4	0.877	2.220	1.027	1.722	1.686	0.028	-0.049	0.021
	1	5	0.927	2.630	1.200	1.812	2.168	0.024	-0.023	0.014
	2	1	-0.380	-2.459	-0.080	-0.181	-0.134	-0.005	-0.015	0.029
	2	2	-0.488	-2.554	-0.030	-0.114	-0.195	-0.004	-0.011	0.019
	2	3	-0.502	-2.460	-0.014	-0.075	-0.174	-0.003	-0.009	0.014
	2	4	-0.572	-2.511	0.013	0.035	-0.216	-0.003	-0.007	0.011
	2	5	-0.622	-2.572	0.052	0.136	-0.270	-0.002	-0.003	0.007
	3	1	0.079	-0.098	-0.008	-0.270	-0.115	-0.217	-0.711	0.215
	3	2	0.078	-0.107	-0.029	-0.246	-0.307	-0.213	-0.682	0.207
	3	3	0.078	-0.075	-0.032	-0.225	-0.332	-0.205	-0.674	0.207
	3	4	0.075	-0.023	-0.037	-0.173	-0.481	-0.199	-0.671	0.210
	3	5	0.072	0.089	-0.146	-0.141	-0.680	-0.201	-0.671	0.218
	4	1	-0.186	0.532	-0.330	-0.010	-0.614	0.021	0.527	-0.260
	4	2	-0.180	0.118	-0.452	0.012	-0.608	0.020	0.481	-0.197
	4	3	-0.181	-0.024	-0.473	0.023	-0.570	0.018	0.438	-0.160
	4	4	-0.173	-0.500	-0.500	0.059	-0.590	0.016	0.393	-0.132
	4	5	-0.167	-1.321	-0.697	0.078	-0.615	0.014	0.276	-0.108
	5	1	-0.137	0.301	-0.368	-0.578	-0.472	0.136	0.064	-0.027
	5	2	-0.128	0.517	-0.385	-0.626	-0.549	0.137	0.061	-0.027
	5	3	-0.129	0.532	-0.383	-0.680	-0.540	0.138	0.062	-0.028
	5	4	-0.122	0.716	-0.375	-0.948	-0.618	0.136	0.062	-0.029
	5	5	-0.120	1.000	-0.547	-1.124	-0.685	0.143	0.068	-0.030

6	1	-0.066	0.008	0.018	-0.543	-0.225	0.014	-0.079	0.207
6	2	-0.076	0.046	0.044	-0.551	-0.335	0.015	-0.082	0.200
6	3	-0.077	0.054	0.048	-0.569	-0.354	0.014	-0.084	0.200
6	4	-0.085	0.098	0.052	-0.694	-0.439	0.013	-0.086	0.205
6	5	-0.090	0.173	0.138	-0.761	-0.578	0.014	-0.095	0.210
Female HH head vs. male HH head									
1	1	-0.265	0.818	-0.665	1.537	-1.944	0.206	-1.328	-0.532
1	2	-0.438	-0.989	-0.517	1.372	-1.698	0.195	-1.175	-0.301
1	3	-0.457	-1.180	-0.472	1.225	-1.657	0.183	-0.960	-0.285
1	4	-0.579	-2.713	-0.395	0.697	-1.528	0.167	-0.811	-0.229
1	5	-0.670	-5.405	-0.531	0.420	-1.389	0.145	-0.386	-0.148
2	1	-1.269	-1.050	-0.015	-1.183	-2.233	-0.033	-1.317	0.334
2	2	-1.125	0.659	-0.291	-1.154	-2.151	-0.026	-1.169	0.213
2	3	-1.123	0.729	-0.351	-1.132	-2.059	-0.023	-1.013	0.161
2	4	-1.012	1.838	-0.446	-1.123	-1.964	-0.020	-0.857	0.130
2	5	-0.934	3.205	-0.731	-1.169	-1.839	-0.013	-0.464	0.079
3	1	0.952	0.835	-0.127	-0.383	1.867	-0.820	5.945	2.043
3	2	1.002	0.676	-0.464	-0.361	1.726	-0.822	5.710	1.970
3	3	1.016	0.573	-0.533	-0.353	1.713	-0.796	5.646	1.971
3	4	1.046	0.288	-0.635	-0.346	1.625	-0.780	5.624	1.993
3	5	1.069	-0.372	-0.889	-0.348	1.528	-0.793	5.660	2.068
4	1	-0.203	-1.080	0.263	0.540	-0.587	-0.038	1.800	0.605
4	2	-0.240	0.097	0.128	0.479	-0.513	-0.034	1.643	0.458
4	3	-0.247	0.427	0.098	0.442	-0.490	-0.034	1.493	0.372
4	4	-0.276	1.997	0.047	0.344	-0.460	-0.030	1.342	0.308
4	5	-0.298	5.590	-0.004	0.286	-0.427	-0.024	0.942	0.251
5	1	1.626	0.665	0.533	0.777	1.401	0.446	-1.609	-0.806
5	2	1.663	-0.211	1.057	0.812	1.265	0.465	-1.568	-0.803
5	3	1.683	-0.308	1.154	0.861	1.225	0.487	-1.599	-0.836
5	4	1.696	-1.075	1.298	1.143	1.169	0.499	-1.630	-0.853
5	5	1.711	-2.445	2.006	1.340	1.063	0.538	-1.805	-0.885
6	1	-0.841	-0.189	0.010	-1.288	1.545	0.226	-2.090	-1.776
6	2	-0.862	-0.232	0.087	-1.149	1.386	0.219	-2.042	-1.712
6	3	-0.872	-0.242	0.104	-1.043	1.327	0.196	-2.063	-1.713
6	4	-0.875	-0.336	0.131	-0.714	1.244	0.178	-2.083	-1.732
6	5	-0.880	-0.573	0.149	-0.530	1.173	0.176	-2.180	-1.798
Primary female is five years older									
1	1	0.463	0.026	0.052	0.365	0.685	-0.050	-0.254	0.120
1	2	0.475	0.058	0.197	0.369	0.784	-0.048	-0.224	0.068
1	3	0.480	0.061	0.226	0.391	0.779	-0.045	-0.183	0.065
1	4	0.484	0.087	0.322	0.546	0.872	-0.042	-0.155	0.049
1	5	0.487	0.127	0.418	0.620	1.008	-0.037	-0.074	0.034
2	1	-0.310	0.138	-0.015	0.100	-0.145	0.010	0.311	-0.022
2	2	-0.313	0.075	-0.072	0.123	-0.160	0.008	0.276	-0.014
2	3	-0.316	0.068	-0.084	0.137	-0.149	0.007	0.239	-0.011
2	4	-0.316	0.030	-0.101	0.182	-0.157	0.006	0.202	-0.009
2	5	-0.317	-0.014	-0.156	0.231	-0.167	0.004	0.109	-0.005
3	1	-0.088	-0.196	0.006	0.033	-0.108	0.118	-0.098	-0.340
3	2	-0.091	-0.158	0.034	0.034	-0.162	0.127	-0.099	-0.328
3	3	-0.092	-0.148	0.040	0.037	-0.170	0.128	-0.098	-0.328
3	4	-0.094	-0.120	0.049	0.051	-0.215	0.130	-0.098	-0.331

3	5	-0.094	-0.085	0.074	0.061	-0.276	0.134	-0.098	-0.344
4	1	0.276	0.162	-0.026	0.053	-0.249	0.005	0.105	-0.185
4	2	0.286	0.109	-0.107	0.061	-0.238	0.004	0.096	-0.140
4	3	0.289	0.094	-0.123	0.064	-0.224	0.004	0.087	-0.114
4	4	0.294	0.046	-0.147	0.080	-0.226	0.004	0.079	-0.094
4	5	0.297	-0.032	-0.228	0.089	-0.229	0.003	0.055	-0.077
5	1	-0.188	-0.071	-0.023	0.060	-0.208	-0.083	-0.340	-0.415
5	2	-0.197	-0.039	-0.084	0.007	-0.226	-0.090	-0.332	-0.413
5	3	-0.200	-0.035	-0.099	-0.040	-0.221	-0.096	-0.339	-0.430
5	4	-0.204	-0.011	-0.114	-0.225	-0.244	-0.100	-0.345	-0.439
5	5	-0.208	0.021	-0.178	-0.344	-0.261	-0.107	-0.383	-0.456
6	1	-0.153	-0.059	0.006	-0.610	0.026	-0.008	0.374	0.927
6	2	-0.159	-0.044	0.032	-0.593	-0.018	-0.010	0.366	0.894
6	3	-0.161	-0.041	0.037	-0.590	-0.027	-0.009	0.369	0.894
6	4	-0.163	-0.031	0.045	-0.633	-0.058	-0.009	0.380	0.904
6	5	-0.165	-0.017	0.071	-0.658	-0.108	-0.011	0.391	0.939
Education of primary female, some vs. none									
1	1	-2.637				-6.732			
1	2	-2.727				-5.833			
1	3	-2.760				-5.690			
1	4	-2.798				-5.210			
1	5	-2.828				-4.685			
2	1	0.493				0.807			
2	2	0.490				0.749			
2	3	0.494				0.723			
2	4	0.488				0.669			
2	5	0.484				0.599			
3	1	0.775				1.632			
3	2	0.806				1.444			
3	3	0.816				1.424			
3	4	0.830				1.296			
3	5	0.840				1.142			
4	1	0.271				0.255			
4	2	0.285				0.189			
4	3	0.289				0.186			
4	4	0.298				0.148			
4	5	0.304				0.106			
5	1	0.392				-0.285			
5	2	0.412				-0.310			
5	3	0.418				-0.303			
5	4	0.428				-0.335			
5	5	0.435				-0.358			
6	1	0.706				1.530			
6	2	0.734				1.332			
6	3	0.743				1.266			
6	4	0.755				1.153			
6	5	0.765				1.035			
Education of primary female, primary vs. none									
1	1		-1.491	-0.116	0.626		0.272	-0.386	0.279
1	2		-0.609	-0.101	0.652		0.255	-0.341	0.158
1	3		-0.481	-0.086	0.722		0.238	-0.278	0.150

1	4	0.144	-0.061	1.144	0.219	-0.235	0.120
1	5	0.977	-0.072	1.367	0.196	-0.112	0.078
2	1	2.033	-0.189	0.516	-0.047	-1.007	0.842
2	2	1.088	-0.987	0.536	-0.037	-0.892	0.538
2	3	0.999	-1.148	0.540	-0.030	-0.772	0.407
2	4	0.434	-1.392	0.570	-0.024	-0.653	0.329
2	5	-0.191	-2.176	0.646	-0.014	-0.353	0.199
3	1	-0.450	0.017	-0.060	-0.763	0.107	0.538
3	2	-0.397	0.081	-0.049	-0.707	0.104	0.519
3	3	-0.349	0.093	-0.034	-0.662	0.103	0.519
3	4	-0.256	0.112	0.015	-0.620	0.102	0.525
3	5	-0.096	0.196	0.049	-0.616	0.104	0.545
4	1	0.313	0.115	-0.098	-0.090	1.261	-0.711
4	2	-0.080	0.272	-0.063	-0.085	1.151	-0.538
4	3	-0.221	0.303	-0.040	-0.078	1.047	-0.437
4	4	-0.698	0.348	0.039	-0.071	0.941	-0.362
4	5	-1.535	0.522	0.090	-0.063	0.661	-0.295
5	1	0.356	0.099	-0.017	0.560	0.490	-1.096
5	2	0.580	0.163	-0.124	0.533	0.478	-1.092
5	3	0.594	0.175	-0.227	0.512	0.487	-1.136
5	4	0.786	0.191	-0.672	0.482	0.497	-1.160
5	5	1.083	0.294	-0.972	0.498	0.550	-1.204
6	1	-0.761	0.115	-0.967	0.110	-0.708	-1.207
6	2	-0.582	0.572	-0.952	0.102	-0.696	-1.164
6	3	-0.542	0.663	-0.960	0.096	-0.702	-1.164
6	4	-0.411	0.802	-1.097	0.089	-0.711	-1.177
6	5	-0.239	1.236	-1.181	0.086	-0.747	-1.222
Education of primary female, secondary vs. none							
1	1	-4.765	-1.067	-1.978	0.387	-4.160	-0.315
1	2	-2.865	-1.058	-1.757	0.363	-3.678	-0.178
1	3	-2.549	-1.033	-1.553	0.338	-3.004	-0.169
1	4	-1.259	-0.981	-0.805	0.309	-2.539	-0.136
1	5	0.361	-1.412	-0.409	0.274	-1.209	-0.088
2	1	4.096	0.066	5.104	-0.072	1.427	1.832
2	2	2.353	-0.089	4.814	-0.057	1.266	1.171
2	3	2.183	-0.128	4.617	-0.046	1.096	0.885
2	4	1.158	-0.192	4.236	-0.037	0.927	0.716
2	5	0.051	-0.362	4.101	-0.022	0.501	0.433
3	1	0.785	-0.197	-0.272	-1.482	5.157	1.211
3	2	0.605	-0.631	-0.245	-1.400	4.956	1.169
3	3	0.635	-0.719	-0.218	-1.319	4.902	1.169
3	4	0.642	-0.850	-0.143	-1.246	4.883	1.181
3	5	0.746	-1.169	-0.092	-1.240	4.919	1.225
4	1	0.848	0.254	-1.249	-0.123	3.564	-1.448
4	2	0.081	-0.231	-1.128	-0.115	3.255	-1.097
4	3	-0.206	-0.332	-1.051	-0.106	2.959	-0.890
4	4	-1.090	-0.494	-0.853	-0.095	2.660	-0.737
4	5	-2.552	-0.874	-0.739	-0.083	1.868	-0.600
5	1	-0.823	0.526	-0.244	1.104	0.931	0.555
5	2	-0.141	0.587	-0.354	1.067	0.908	0.551
5	3	-0.052	0.590	-0.465	1.036	0.926	0.574

5	4	0.454	0.588	-0.965	0.986	0.945	0.588
5	5	1.141	0.917	-1.304	1.019	0.963	0.609
6	1	-0.141	0.417	-1.362	0.205	-6.439	-2.950
6	2	-0.034	1.422	-1.330	0.193	-6.309	-2.843
6	3	-0.011	1.675	-1.329	0.180	-6.362	-2.845
6	4	0.094	1.927	-1.470	0.168	-6.440	-2.880
6	5	0.253	2.900	-1.557	0.165	-6.757	-2.989
Education of primary female, secondary vs. primary							
1	1	-3.274	-0.909	-2.604	0.115	-3.774	-0.594
1	2	-2.256	-0.957	-2.409	0.108	-3.337	-0.336
1	3	-2.068	-0.947	-2.275	0.100	-2.725	-0.319
1	4	-1.403	-0.919	-1.949	0.091	-2.303	-0.256
1	5	-0.616	-1.340	-1.776	0.079	-1.097	-0.166
2	1	2.063	0.255	4.588	-0.025	2.434	0.990
2	2	1.265	0.898	4.278	-0.020	2.158	0.633
2	3	1.185	1.020	4.077	-0.016	1.868	0.478
2	4	0.724	1.201	3.665	-0.013	1.580	0.387
2	5	0.241	1.814	3.454	-0.008	0.854	0.296
3	1	1.235	-0.214	-0.212	-0.719	5.050	0.673
3	2	1.002	-0.712	-0.196	-0.693	4.853	0.650
3	3	0.984	-0.812	-0.184	-0.657	4.799	0.650
3	4	0.897	-0.962	-0.157	-0.626	4.781	0.656
3	5	0.841	-1.365	-0.142	-0.624	4.815	0.681
4	1	0.535	0.139	-1.151	-0.033	2.304	-0.664
4	2	0.162	-0.504	-1.066	-0.030	2.108	-0.558
4	3	0.015	-0.635	-1.011	-0.028	1.912	-0.453
4	4	-0.392	-0.841	-0.893	-0.024	1.719	-0.375
4	5	-1.017	-1.396	-0.829	-0.020	1.207	-0.305
5	1	-1.179	0.427	-0.226	0.543	0.441	1.652
5	2	-0.721	0.423	-0.230	0.534	0.430	1.642
5	3	-0.646	0.415	-0.238	0.524	0.439	1.710
5	4	-0.332	0.397	-0.293	0.504	0.448	1.748
5	5	0.059	0.623	-0.332	0.521	0.497	1.813
6	1	0.621	0.302	-0.395	0.094	-5.730	-1.743
6	2	0.548	0.851	-0.377	0.091	-5.880	-1.680
6	3	0.531	0.959	-0.369	0.084	-5.661	-1.681
6	4	0.504	1.311	-0.373	0.079	-5.729	-1.703
6	5	0.492	1.663	-0.376	0.080	-6.010	-1.767

Chapter 2

Returns to Public Expenditures on Agricultural Research and Extension in Virginia: An Attempt to Solve the Ad Hoc Model Selection

Most decisions about public investments in agricultural research and extension (R&E) are made at the state level. In Virginia, about 70% of the funds to Virginia Cooperative Extension/Agriculture Experiment Station (Agency 229) from 1996-2010 come from the state general fund (Fornash, 2011). By contrast, most evaluations of aggregate R&E benefits are made on a regional or national basis. We are aware of only a few state-level evaluations, e.g., Norton, Coffey and Frye (1984), Norton and Paczkowski (1993), Alston, Pardey and Carter (1994), Alston, et al. (2010). Understanding the relationship between agricultural productivity and R&E at the state level enables state policymakers to make more informed decisions about resource allocations, agricultural policies, employment generation, and rural development through investments in agriculture.

Evaluations of the aggregate benefits of agricultural R&E at national and state levels follow the dame rules, but some subtle differences exist in applications. For instance, Acquaye, Alston and Pardey (2003) argue that using pre-aggregated data biases national-level input and output indices. Ball, et al. (1999) emphasize that, while interstate transactions in intermediate inputs must be recognized for creating state production accounts, these transfers must be ignored for national accounts. Spillover effects, usually ignored for national-level evaluations, must be considered in state-level evaluations.

The primary objective of this study is to evaluate the returns to public expenditures on agricultural R&E in Virginia using the latest data. An equivalently important objective is to consider the model uncertainty in these kinds of assessments. Model uncertainty arises because researchers use different control variables, functional forms, and most importantly, different distributed lags for generating knowledge stocks in the context of evaluating agricultural R&E. Previous studies offer quite uncertain rate-of-return estimates. Alston, et al. (2000) collect 1,852 effective rate-of-return estimates from 292 studies in 1953-1998. Estimated rates of return from studies focusing on agriculture alone span a wide range: -7.4% to 5,645% for research and 0% to 636% for extension. The distribution is also flat: only 21% of the estimates fall within 40-60%, the often-cited "normal" range. Because the studies focus on various regions and periods, utilize data from different sources, and adopt different methods of evaluation, the wide range of rates of return is not surprising. Besides these factors, model uncertainty is undoubtedly an important source of uncertain results.

The econometric modeling strategies in the literature can be classified into six categories that will be discussed below. After comparing their pros and cons, we conclude that Bayesian Model Averaging (BMA) and Bayesian Hierarchical Model (BHM) are promising methods to address model uncertainty. Model averaging is particularly important for policy evaluations. Policymakers do not want to condition their decisions on a single model unless the model is true with a high degree of certainty (Brock, Durlauf and West, 2003). The results given by model averaging involves both the uncertainty of models and the uncertainty of coefficients given a model, reduce the risk of conditioning decisions on an inappropriate model, and offer the a margin of error for policymakers. Although BMA and BHM reduce arbitrariness compared with the ad hoc choice of lag structures in most studies, arbitrariness has not been eradicated: BMA depends on a pre-determined set of lags while BHM depends on an assumed functional form. When the BMA model space is filtered with tests of statistical adequacy, a particular functional form emerges that could be applied in BHM. On the other hand, BHM directly estimates the lag structures and serves as a remedy to the remaining arbitrariness in BMA, which offers the incentive to combine the two methods.

By BMA analysis, the spatial medians of the internal rate of return (IRR) of Virginia's agricultural R&E are 35.8% and 28.8% for research and extension respectively, and 20.6% of the agricultural productivity growth from 1949 to 2016 comes from own investments in agricultural R&E. By BHM, the spatial medians are 26.4% and 42.4%, and own investments contribute 19.0% of the productivity growth. The results by BHM are preferable because the lag structures are less restricted in BHM estimation and the estimated lag shapes are more reasonable than those by BMA. We also calculate the modified IRR (Lin, 1976) and obtain the distributions of the rate-of-return estimates, which is usually ignored in the literature.

The rest of the paper is structured as follows. We first review the econometric strategies used in evaluations of agricultural R&E, raise the issue of model uncertainty, and introduce the essentials of BMA and BHM. After a brief introduction of the Gamma distributed lag and tests of statistical adequacy, the details about implementing BMA and BHM, calculating elasticities and rates of return, and obtaining their distributions are presented. Then we present the dat and discuss the results of econometric analysis. Finally, we conclude the study, focusing on revisiting the methodology.

2.1 Econometric strategies of evaluating the returns to agricultural R&E

Except for a few exceptions (e.g., Andersen, 2015), ex-post econometric assessments of agricultural R&E obtain their results from models that fall into two categories: statistical meta-function methods and statistical productivity decomposition methods. The former includes R&E directly in a production, profit, or cost function, while the latter begins with a productivity index and relates it to R&E expenditures (Huffman and Evenson, 2006). We adopt the second approach because it avoids the specification of a particular functional form for the production or profit function and the possible endogeneity of inputs.

An agricultural productivity index measures the changes in agricultural outputs that can not be explained by the growth of conventional inputs, i.e., the *neglected inputs* (Schultz, 1956). Agricultural technology advancements, as the products of agricultural R&E, account for most productivity growth. Explaining the index directly with lagged R&E expenditures is impractical because of serial correlation of expenditures. Since the effects of R&E are spread over many years, it is necessary to include many lags and, as a result, the model is unlikely to produce meaningful coefficient estimates. Multicollinearity is a major issue in the econometric analysis of the linkage between productivity and agricultural R&E (Griliches, 1979). Another problem of using lagged expenditures as covariates is the oscillating signs of estimated coefficients of successive lags, which is difficult to rationalize.

In order to address serial correlation, researchers assume that technology is the result of a *research production function* that takes the form of a weighted average of lagged expenditures (Evenson, 1967):

$$S_t = \sum_{j=0}^{T-1} w_j I_{t-j}\, , \sum_{j=0}^{T-1} w_j = 1 \tag{2.1}$$

Knowledge stock (S_t), the average of lagged expenditures I_{t-j} ($j = 0,1,\dots,T-1$), is hypothesized to improve agricultural productivity in a similar fashion as physical capital. The weights for calculating the knowledge stocks, or *lag weights*[15] (w_j, $j = 0,1,\dots,T-1$), are assumed to follow a pattern that mimics the lifespan of flows of impacts from an ordinary agricultural research project. A project takes several years to complete (gestation). After completion of a successful project, farmers gradually adopt its outcomes. The adoption rate climbs, reaches a peak, and falls as new technologies emerge to replace the outdated (disadoption).

Although agricultural economists have reached a broad agreement about the general shape of the lag structure (lag weight increases and then decreases with j), its precise shape is far from clear. Various mathematical forms can represent the general shape, e.g., inverted-V (Evenson, 1967), trapezoidal (Huffman and Evenson, 2006; Jin and Huffman, 2016), Almon polynomial (Yee, 1992; Norton, Coffey and Frye, 1984; Norton and Paczkowski, 1993), Beta distribution (Thirtle, Piesse and Schimmelpfennig, 2008), and Gamma distribution (Alston, et al., 2010; Baldos et al., 2018). By Alston, et al. (2000), the most used mathematical forms in the literature are polynomial (15.4%), trapezoidal (4.1%), and inverted-V (2.7%). The length of lags (T) also varies: 0-10 years for 14.7%, 11-20 years for 16.4%, 21-20 years for 5.1%, and 31-40 years for 4.5% of the studies[16]. Recent studies tend to use longer lags. Pardey and Craig (1989) argue for using lags at least 30 years to capture the complete impacts of research based on causality tests. Alston, et al. (2010) assume a 50-year lag structure for US agricultural research, and the same lag is used by Andersen and Song (2013). Jin and Huffman (2016) use a 35-year lag structure.

The shape and length of the lag structure can have a sizable influence on rate-of-return estimates. However, the method of determining the lag structures is usually unclear in evaluations of agricultural R&E and other studies that also employ distributed lags[17]. Six approaches are used

[15] In this article, a particular time series of lag weights is defined as a lag structure, which is characterized by its length and shape.
[16] The percentages by mathematical forms/lag length do not add up to one because about 2/3 of the studies do not clearly describe their lags.
[17] A large number of epidemiological studies employ distributed lag models to inspect the relationship between environmental factors and the occurrence of certain diseases or death rate, for instance, the relationship between air pollution and mortality rate (Welty, et al., 2009) and the relationship between heat waves and mortality rate (Bobb, Dominici and Peng, 2011). Their problem is similar to ours: the number of patients in a day is related to the severity of pollution or temperature in previous days in some unknown pattern.

to determine the lag structures in previous studies. 1) Lag structures are selected ad hoc or from those used in other studies (Plastina and Fulginiti, 2012; Andersen and Song, 2013; Jin and Huffman, 2016). 2) Lag structures are selected from a pre-determined set by criteria such as R^2 or information criterion (Evenson, 1967; Alston, et al., 2010). 3) A mathematical form is assumed for the lag structure, and the parameters of the mathematical form are estimated by maximum likelihood (Thirtle, Piesse and Schimmelpfennig, 2008). 4) A group of models based on different lag structures is averaged using model averaging (Balcombe and Rapsomanikis, 2010; Bobb, Dominici and Peng, 2011). 5) Some loose constraints instead of particular mathematical forms are assumed for the lag weights, and the lag weights are estimated directly (Shiller, 1973; Welty, et al., 2009). 6) A productivity index is related to knowledge stocks in a *top model* while the knowledge stock is related to lagged expenditures in a *bottom model*; the system is estimated using the Bayesian Hierarchical Model (Baldos, et al., 2018).

Approach 1, adopted by the majority of the studies for ease of use, is not preferred because of its arbitrary choice of lag structures. The problem with Approach 2 is the pre-determined set may not include the "correct" lag structure. Besides, selecting lag structures by Goodness-of-fit or the information criterion is questionable. Goodness-of-fit is purely a mathematical concept about how well a curve fits the data points. Information criteria are transformations of likelihood that measures the possibility that a data set arises from the model. Neither of them concerns about whether the statistical assumptions are correctly specified or the selected lag structure is reasonable. Since only a single lag structure is selected at the end, the uncertainty of lag structures is still neglected.

Estimating lag structures directly by ML (Approach 3) is valid but constrained by two issues. First, convergence is a challenge. We attempt to estimate a Gamma distributed lag with a specification of linear model and autoregressive errors using ML, and convergence cannot be attained unless the criterion of convergence is relaxed[18]. The more serious problem is that, even if the estimation converges, obtaining the distribution of lag structure parameters is difficult. Thirtle, Piesse and Schimmelpfennig (2008) try Gamma, Beta, and polynomial distributed lags and estimate the parameters of lag structures directly with UK data. The authors report the estimated lag structure but ignore its uncertainty. The appendix of the paper shows they have obtained the

[18] Estimation converged when the relative change in parameters falls below 0.01% between two iterations, which is considerably larger than the criterion commonly used in ML estimation.

standard errors of the lag parameters. However, the standard errors cannot be correct because the parameters have to be positive and a constrained ML estimation is required. This is possibly the reason why the authors only report the "point" estimates of lag structures. A large sample allows using bootstrapping to obtain the distribution of the lag parameters. However, the data set for evaluating returns to R&E is usually a small size (<100), so bootstrapping is not applicable.

The advantage of Approach 5 is it only imposes minimal prior knowledge on the lag weights and avoids assuming an inappropriate mathematical representation. For instance, Shiller (1973) assumes the lag structure to be smooth: $w_{j-1} + w_{j+1} - 2w_j \sim \mathrm{N}(0, \xi)$. Welty, et al. (2009) assume the influence of air pollution on mortality rate diminishes to zero as the lag increases: $\lim_{j\to\infty} w_j = 0$. However, unlike a particular mathematical representation that reduces the number of unknown parameters to one or two, implementing Approach 5 still needs to estimate many parameters. Therefore, it is applicable only when the lag length is short or a large sample is available, neither of which is true for this study. We believe Approaches 4 (model averaging) and 6 (BHM) are feasible and briefly introduce them in the next section, though each of them has its advantages and weaknesses.

2.2 A brief introduction of BMA and BHM

Choice of explanatory variables is the main source of model uncertainty if the dependent variable is fixed. Approaches commonly used in variable selection include stepwise regression (forward, backward), shrinkage methods (ridge regression, LASSO regression), information criteria (AIC, BIC), and cross-validation. However, all these methods are subject to the *model selection bias* that results from selecting a model and presenting it as the "true" model based on the same set of data (Chatfield, 1995). This is problematic because any data set embodies some patterns that might be only occasional. The "true" model should capture the patterns in multiple data sets, and thus should be established using multiple data sets. These data-driven model selection approaches are usually called *data mining* (or *data fishing*, *data dredging*) and believed to be "logically unsound and practically misleading" (Zhang, 1992).

Instead of these approaches, we consider two criteria for judging a model: statistical adequacy (SA) and prior (subject) knowledge about the relationship between agriculture R&E and productivity. The former means that the underlying statistical assumptions of a model are tested and justified, while the latter means that the model is formulated based on a proper economic theory and its results can be easily rationalized. We will revisit the criteria and discuss implications after presenting the results.

None of the criteria could select a single "true" model for us; they only identify a group of models, each of which is defensible. The idea of combining estimated coefficients from different models dates back to Laplace in 1818 (Clemen, 1989). Bayesian approaches for combining estimates emerged in the 1960s and 1970s, e.g., Geisser (1965), Geisel (1973). However, Bayesian approaches have only been widely adopted since the 1990s, when powerful computers and convenient methods of implementation (e.g., Markov Chain Monte Carol Model Composition, MC^3) emerged.

As its name suggests, the method of Bayesian Model Averaging (BMA) originates from Bayes' theorem. Given H different models M_1, M_2, …, M_H, by Bayes' theorem, the posterior probability of M_i, $i \in \{1,2,\dots,H\}$ given data D is

$$P(M_i|D) = \frac{P(D|M_i)P(M_i)}{P(D)} \tag{2.2}$$

$P(M_i)$ is the prior probability of M_i, and $P(D)$ is the likelihood by all H models:

$$P(D) = \sum_{i=1}^{H} P(D|M_i)P(M_i) \tag{2.3}$$

Suppose Δ is some quantity of interest (e.g., rates of return of agricultural R&E) and Δ_i is the estimated quantity based on M_i, the model-averaged Δ is

$$\overline{\Delta} = \sum_{i=1}^{H} \Delta_i P(M_i|D) \tag{2.4}$$

The difficulty of implementing BMA is calculating the model likelihood $P(D|M_i)$ because it involves the integration over the parameter space given model. Raftery (1995) derives an approximation of the model likelihood and relates it to the Bayesian Information Criterion (BIC). Sala-i-Martin, Doppelhofer and Miller (2004) apply the approximation to linear models with normal independent identically-distributed (NIID) error terms and name their approach Bayesian Averaging of Classical Estimates (BACE). Because it uses information criterion to average models, which is similar to the method of Frequentist Model Averaging (FMA)[19], this method is called the *frequentist BMA* by Moral-Benito (2015). Compared with the common BMA, BACE is easier to implement because the prior distribution of parameters is not needed, and only model priors need to be specified. This reduces computation load, and more importantly, allows a combination of statistical adequacy tests with model averaging. Details about BACE are presented in Appendix 2.1.

Although the original BACE method focuses on OLS and variable selection, it does not need to be restricted. Bryant and Davis (2008) extend the original method to systems of equations, endogenous covariates, and various functional forms. Moral-Benito (2012) extends the approach to a panel data framework to address endogeneity and model uncertainty simultaneously and names the method Bayesian Averaging of Maximum Likelihood Estimates (BAMLE). We are aware of few studies that adopt BMA in evaluating agricultural R&E. Balcombe and Rapsomanikis (2010) apply BMA to a model space that considers various lag structures and time of structural change. However, their choice of allowing knowledge stocks based on different lag structures to enter a model simultaneously is unusual.

Instead of BMA, Baldos, et al. (2018) explicitly estimate the parameters of a 50-year Gamma distributed lag using a Bayesian Hierarchical Model (BHM)[20] . The key characteristic of a hierarchical model is the knowledge stocks are taken as unknown parameters determined by other data and parameters. The *top model* links the productivity index with knowledge stocks and other control variables, and the *bottom model* relates knowledge stocks to R&E expenditures. We adopt this two-stage specification but extend it on several aspects. Baldos, et al. (2018) evaluate the returns to agricultural research in the U.S. at the national level, consider only research, and ignore spillovers. The state-level evaluation in this study includes both research and extension and considers the spillover effects.

BHM avoids determining a set of lag structures beforehand and estimates the lag structures directly, but it requires the specification of the model functional form at the beginning, which is subject to the same arbitrariness in any empirical study that concludes from a single model. For instance, Baldos, et al. (2018) use a log-log model and a corn moisture index as the proxy for the

[19] FMA differs from BACE in that the weights for averaging estimates in FMA do not involve the prior probabilities.

[20] BHM is more often used in epidemiology. For example, Huang, Dominici and Bell (2005) use a two-stage model to inspect the association between ozone levels and cardio-respiratory mortality. Their bottom model is different: they model the total effects (sum of weights) instead of individual lag weights.

weather but do not explain the choices. The model space in a BMA analysis can include models of various functional forms and control variables. For explaining the relationship between heat waves and mortality rate, Bobb, Dominici and Peng (2011) consider various measures of temperature (daily maximum/average) and interaction of temperature and humidity. We include models with various functional forms into the model space of BMA analysis. As we show below, when we screen the model space of statistical adequacy tests, a particular functional form emerges and can be used for BHM estimation.

2.3 Gamma distributed lag

Jorgenson (1966) suggests two criteria for a proper distributed lag. First, it should be possible to approximate an arbitrary lag to any desired degree of accuracy. Second, the number of unknown parameters should be as small as possible. Therefore, among the various mathematical representations, we choose the Gamma distributed lag for its ability to mimic the shapes of other lag types, and only two parameters are required to define its structure. The lag weight of year j ($j = 0, 1, \ldots, T-1$), given shape parameter a, rate parameter b, and lag length T, is:

$$w_j = \frac{(j+1)^{a-1}e^{-b(j+1)}}{\sum_{i=0}^{T-1}(i+1)^{a-1}e^{-b(i+1)}}, a > 0, b > 0 \tag{2.5}$$

The shape of the lag structure tends to be flat for small b and steep for large b. It is easy to prove that w_j is maximized and a lag structure peaks at $j = (a-1)/b - 1$, so given b we could vary a to change the peak of the lag. Therefore, by finely tuning a and b, the shapes of many lag structures can be mimicked. A steep Gamma distributed lag can be taken as triangular, while a flat one is similar to a trapezoidal lag. The lag weights of the first few years could be reduced to near zero to mimic the gestation period. Similar steps allow the last few years to conform to a lag structure shorter than T.

2.4 Tests of statistical adequacy (SA)

Statistical models depend on assumptions, which need to be tested to make correct inferences (Mayo and Spanos, 2004). In this study, hypotheses about the effects of agricultural R&E on productivity are based on the model that relates productivity to R&E investments. The validity of such tests rests on testable assumptions about the model. Models showing negative impacts can be eliminated because it is highly unlikely to be correct. Significance tests of such negative impacts are premised on the idea that the underlying model assumptions are met.

Classical linear regression models (CLRMs) are adopted in this study. For testing whether a CLRM is statistically adequate, we follow the comprehensive approach suggested by McGuirk, Driscoll and Alwang (1993). The approach includes three tests: a test of normality, a joint test of misspecification (conditional mean), and a joint test of conditional variance. For testing the normality of estimated residuals, we adopt the D'Agostino-Pearson omnibus test that combines the tests of skewness and kurtosis (D'agostino, Belanger and D'Agostino Jr, 1990). The joint tests of conditional mean and conditional variance are based on the following auxiliary regressions:

$$\hat{u}_t = \boldsymbol{\beta}'_0\boldsymbol{X}_t + \boldsymbol{\Gamma}'_P\boldsymbol{\Psi}^P_t + \boldsymbol{\Gamma}'_F\boldsymbol{\Psi}^F_t + \boldsymbol{\Gamma}'_I\boldsymbol{\Psi}^I_t + v_t \tag{2.6}$$

$$\hat{u}^2_t = \alpha + \boldsymbol{\Gamma}'_P\boldsymbol{\Psi}^P_t + \boldsymbol{\Gamma}'_S\boldsymbol{\Psi}^S_t + \boldsymbol{\Gamma}'_D\boldsymbol{\Psi}^D_t + v_t \tag{2.7}$$

$\hat{u}_t$ and $\boldsymbol{X}_t$ are predicted residuals and the matrix of covariates of the main regression. $\boldsymbol{\Psi}^P_t$ is a binary variable that distinguishes the first and second halves of the sample for testing structural change. $\boldsymbol{\Psi}^F_t = \boldsymbol{\Psi}^S_t = (\hat{y}^2_t,\ \hat{y}^3_t)$ is for testing functional form in conditional mean test and static heteroscedasticity in conditional variance test ($\hat{y}_t$ is predicted dependent variable). $\boldsymbol{\Psi}^I_t = \hat{u}_{t-1}$ is for testing independence. $\boldsymbol{\Psi}^D_t = \hat{u}^2_{t-1}$ is for testing dynamic heteroscedasticity. The null hypotheses are tested by *F*-tests:

$$\boldsymbol{\Gamma}^P = \boldsymbol{\Gamma}^F = \boldsymbol{\Gamma}^I = \boldsymbol{0} \tag{2.8}$$

$$\boldsymbol{\Gamma}^P = \boldsymbol{\Gamma}^S = \boldsymbol{\Gamma}^D = \boldsymbol{0} \tag{2.9}$$

If the null hypotheses cannot be rejected, the assumptions underlying CLRMs cannot be rejected thus believed to be SA. Since the consequences of model misspecification are severe, McGuirk, Driscoll and Alwang (1993) suggest using a *p*-value of 0.2 to 0.25 as the threshold for tests, and we choose 0.2 for all the three tests in this study.

2.5 Implementation of BMA

Implementation of BMA and BACE begins with the definition of the model space and determining model prior probabilities. It is not possible to estimate all the models in the model space because of its huge size. Therefore, we show the estimation procedure, focusing on the method of stratified sampling to accelerate convergence.

Model space

Since the dependent variable is fixed in all models, defining the model space is reduced to choosing the potential covariates. Virginia's own agricultural R&E knowledge stocks based on various lag structures are added to each model. The research knowledge stock in other states is added to evaluate spillover effects. Due to differences in agro-ecological conditions and agricultural outputs, research conducted in some states is not relevant for Virginia's agricultural productivity. As a result, spill-in research knowledge stocks are generated based on various lag structures and sets of states. Due to data availability, private-sector research is ignored. Since research expenditures and knowledge stocks grow with time, time trends are added to capture the effects of neglected research[21]. Besides technology, another potential source for the variation of agricultural productivity is climate, which significantly influences yield but ignored when calculating the productivity index. We consider two variables to proxy climate: departures of precipitation and the Palmer drought severity index (PDSI) from their normal levels in July[22].

As many models as possible should be included to avoid omission of important models, but larger model spaces take more time to estimate. Various functional forms were tested. Only log-linear models passed the SA tests, so model space is reduced. The final model space includes CLRM represented as:

$$\begin{aligned}\ln MFP_t = {} & \beta_0 + \beta_1(OWN_t + EXT_t) + \beta_2 SPILL_t + \beta_3 CLIM_t \\ & + \beta_4(OWN_t + EXT_t)^2 + \beta_5 SPILL_t^2 + \beta_6(OWN_t + EXT_t) \\ & \times SPILL_t + \beta_7(OWN_t + EXT_t) \times CLIM_t + \beta_8 SPILL_t \times CLIM_t \\ & + \beta_9 t + \beta_{10} t^2 + \varepsilon_t, u_t \sim \mathrm{NIID}(0, \sigma^2)\end{aligned} \tag{2.10}$$

OWN, EXT, and $SPILL$ are own research, own extension, and spill-in research knowledge stocks, respectively. $CLIM$ is climate condition. For coding convenience, a model is represented by a vector of ten elements $\boldsymbol{\gamma} = \{\gamma_1, \gamma_2, \dots, \gamma_{10}\}$. The values that each element could take and their definitions are explained in Table 2.1. Compared with a common model averaging task that allows any combination of covariates, the model space in this study is restricted. OWN,

[21] Adding time trends as covariates is equivalent to running regressions on detrended time series (Greene, 2012). The time series of productivity index and knowledge stocks are clearly trended and non-stationary. As the result, we find models without time trends never pass the SA tests.

[22] We have tried four different drought indices: PDSI, Palmer hydrological drought index (PHDI), Palmer modified drought index (PMDI), and Palmer Z-index (PZI). Since they produce similar results, only PDSI in considered when generating the model space.

EXT, and $SPILL$ are added to each model. t is also added to each model because models without it are never SA. $CLIM$, t^2, and the quadratic terms and interactions of $OWN + EXT$, $SPILL$, and $CLIM$ are all optional. $CLIM^2$ is ignored because it is seldom included in SA models. In addition, we assume that own R&E knowledge stocks together contribute to Virginia's agricultural productivity and combine OWN and EXT in Equation 2.10. Otherwise, obtaining meaningful estimations is difficult because of the strong correlation between OWN and $SPILL$.

For generating the set of lag structures, we fix $T = 25$ for research and $T = 9$ for extension, the maximum lengths allowed by the time series[23]. Twelve values for the b parameter (0.2, 0.4, …, 2.4) are considered. The peak of the lag structure is allowed to vary between year 3 to 24 for research and between 3 to 8 for extension. By the relationship between the peak and a and b parameters, we could obtain a for each (b, peak) pair and calculate the lag weights. That offers 264 (12×22) research lags and 72 (12×6) extension lags, and own research, own extension, and spill-in research knowledge stocks are generated on these lag structures, which cover a wide variety of lag shapes[24].

When generating the spill-in research knowledge stock, since some states might be irrelevant for Virginia, we consider the research in three sets of states: a) the five neighbor states of Virginia, b) 21 states that are agro-ecologically similar to Virginia (see Table 2.1 for the list of states), and c) all states except for Virginia. To measure the similarity between states, we employ EPA's Ecoregions (Omernik and Griffith, 2014) that defines an ecoregion as an area where ecosystem and natural resources are generally homogenous. We use the formula of Alston, et al. (2010) to calculate a similarity index between states but replace the mix of agricultural outputs with the mix of ecoregions. The similarity index is between zero and one; zero means two states do not have any overlap, while one means two states overlap exactly. We select the 21 states for which the similarity index between them and Virginia is greater than 0.1. The similarity indices between Virginia and the unselected states are almost zero.

Model prior probabilities

An easy way of assigning prior probabilities to models is to assume diffuse priors $P(M_i) = P(M_j)$, $\forall\ i \neq j$, reflecting no prior knowledge of model appropriateness. Considering that

[23] Recent studies tend to use longer lags (>30 years), though the lag length does not show a clear evolution with time (Alston, 2010). Our lag length is restricted by data availability.
[24] Appendix Figures 2.1 and 2.2 show a subset of these lags.

knowledge stocks based on different lag structures are highly correlated, we adopt the dilution priors suggested by George (2010) instead of diffuse priors. We first employ the *conditional independence* and *inheritance* assumptions by Chipman (1996) to simplify the calculation of priors. The two assumptions reduce determining the model priors to assigning prior inclusion probabilities to knowledge stocks. We could assign the same probability to each knowledge stock[25]. However, knowledge stocks are highly correlated, and some of them have more similar peers than others do because the set of alternative lag structures are from a $b \times$ peak grid. A lag structure in the center of the grid has more neighbors than those at the corner or edge. Besides, a spill-in research stock is more similar to a stock generated on the same set of states than to one based on a different set. We need to *dilute* the prior inclusion probabilities of stocks with higher similarity to other stocks, following the same logic underlying the car/red bus/blue bus metaphor that challenges the Independence of Irrelevant Alternatives (IIA) assumption. Details about the model priors are presented in Appendix 2.1.

Estimation procedure

Figure 2.1 presents the overall workflow of implementing BACE; the left and right parts show the steps of estimation and postestimation, respectively. Estimating all models and calculating Δ is computationally demanding and impractical because of the huge model space. Sala-i-Martin, Doppelhofer and Miller (2004) suggest randomly drawing models from the model space until the sampling distribution of Δ approaches the true distribution. In practice, the authors check convergence when receiving every additional 10,000 draws, and the estimation is believed to be converged when the changes in the posterior means of OLS coefficients fall below 1×10^{-6} and the changes in the posterior inclusion probabilities of covariates (the fraction of draws that include the covariate) fall below 1×10^{-4}.

Our procedure is essentially the same as that used by Sala-i-Martin, Doppelhofer and Miller (2004), with an extra step of testing the SA of each model drawn from model space. We also adopt their *stratified sampling* method for accelerating convergence. If each model is drawn with the same probability, on average, only one SA model is obtained per 10,000 draws, and convergence is slow. Therefore, diffuse sampling weights are used for initial draws to reduce the error of

[25] For simplicity, a time series of own research or own extension knowledge stocks based on a lag structure is referred to as *a knowledge stock*. For spill-in research, a knowledge stock is generated on a lag structure and a set of states.

sampling, and posterior inclusion probabilities are used as sampling weights afterward. By doing that, model subspaces that are unlikely to pass SA tests are visited less frequently, allowing more rapid convergence.

Specifically, diffuse sampling weights are used until 10,000 SA draws are obtained. Posterior inclusion probabilities are subsequently used. Convergence is checked using posterior covariate inclusion probabilities following attainment of an additional 10,000 SA draws. Convergence is attained when the changes in posterior inclusion probabilities of knowledge stocks fall below 1×10^{-5} and the changes in climate variables, time trends, quadratic terms, and interactions fall below 1×10^{-3}. A stricter criterion is used for knowledge stocks because of the number of alternative knowledge stocks; the inclusion probability of each of them is small. Convergence is checked only by posterior inclusion probabilities because checking the convergence of coefficients is unnecessary. As shown below, the number of SA models are limited, and they are concentrated in a small model subspace. As the posterior inclusion probabilities converge, almost all the SA models have been extracted from the model space.

2.6 The implementation of BHM

Estimating a BHM requires specification of the model functional form and parameter priors. The functional form is obtained when the model space of BMA is filtered with SA tests. Since analytical solutions of posterior probabilities are unavailable, Gibbs Sampling and Metropolis-Hastings algorithm are employed to obtain a simulated sample of parameters.

Model specification

As shown in the section of results, SA tests in BMA estimation select the following *top model* that links productivity with knowledge stocks[26]:

$$\begin{aligned} \ln MFP_t = \beta_0 + \beta_1(OWN_t + EXT_t) + \beta_2 SPILL_t + \beta_4(OWN_t + EXT_t)^2 \\ + \beta_5 SPILL_t^2 + \beta_6(OWN_t + EXT_t) \times SPILL_t + \beta_9 t + \varepsilon_t \end{aligned} \tag{2.11}$$

The *bottom models* relate knowledge stocks to R&E expenditures:

[26] Equation 2.11 uses the same subscripts for coefficients as Equation 2.10.

$$OWN_t = \sum_{j=0}^{24} w_{OWN,j} I_{OWN,t-j} + \varepsilon_{OWN,t} \qquad (2.12a)$$

$$EXT_t = \sum_{j=0}^{8} w_{EXT,j} I_{EXT,t-j} + \varepsilon_{EXT,t} \qquad (2.12b)$$

$$SPILL_t = \sum_{j=0}^{24} w_{SPILL,j} I_{SPILL,t-j} + \varepsilon_{SPILL,t} \qquad (2.12c)$$

Error terms in the bottom models imply that the relationship between knowledge stocks and R&E expenditures is not necessarily deterministic, which is a more realistic assumption. All error terms are assumed to be NIID: $\boldsymbol{\varepsilon} \sim \mathrm{N}(\mathbf{0}, \sigma^2\mathbf{I})$[27]. The variance of the error terms σ^2 is allowed to vary with models.

Priors of parameters

Common conjugate priors are adopted for $\boldsymbol{\beta} = \{\beta_0, \beta_1, \beta_2, \beta_4, \beta_5, \beta_6, \beta_9\}$ and the variance of error terms: normal distribution for $\boldsymbol{\beta}$ and inverse Gamma distribution for σ^2's. The conjugate priors make the conditional posterior distributions also normal and inverse Gamma. Because the lag structure parameters have to be positive, their prior distributions are assumed to be truncated normal with lower bound of zero. The means of the truncated normal distributions are the lag parameters with the largest posterior inclusion probabilities in BMA. To minimize the influence of priors, we assign a large variance (100) to the prior distribution of each parameter. Table 2.2 presents details of the prior distributions.

Gibbs sampling, Metropolis-Hastings algorithm, and criteria of convergence

Since the analytical solution to the joint posterior distributions of parameters is unavailable, we have to resort to the method of Gibbs Sampling to obtain a simulated approximation. Gibbs Sampling (GS), a method of Markov Chain Monte Carol (MCMC) simulation, allows us to obtain a random sample from a probability distribution when the joint density is unavailable, but calculating marginal densities is possible (Casella and George, 1992). For implementing GS, the set of unknown parameters is divided into several subsets, and the posterior probability of each

[27] It is possible that some common missing factors influence the formation of research and extension knowledge stocks. Therefore, it is more reasonable to assume the error terms of different models are dependent. Since that will bring additional complexity to the estimation, we decide to leave it to future work.

subset conditional on other subsets are calculated. After that, simulated parameters are sequentially from the conditional posterior probability of each subset.

For knowledge stocks and lag structure parameters, the conditional posteriors do not follow any known probability distribution. We adopt the Metropolis-Hastings (M-H) algorithm, a generalization of GS, to obtain draws from the distributions (Metropolis, et al., 1953; Hastings, 1970). When directly drawing from the conditional posterior probabilities is impossible, M-H uses a method that is similar to the acceptance-rejection (A-R) sampling to obtain draws: we first draw a parameter from a proposal distribution and then determine whether accepting the draw or not. Since GS and M-H are mixed, the approach can be called *M-H within Gibbs sampling* (Chib and Greenberg, 1995). Details about dividing the parameter set, obtaining conditional posterior densities, and implementing the M-H algorithm are presented in Appendix 2.2.

M-H can be classified into two types by the proposal distribution. The *random-walk chain* obtains a new draw based on the previous draw, while the *independent chain* obtains a new draw independent of the previous draw. We adopt the random-walk chain and use truncated normal distribution as the proposal distribution (Table 2.2). For the b parameters of lag structures, the lower and upper bounds of proposal distributions are identical to those used for defining the BMA model space. However, the bounds of the a parameters are set so that the peaks of lags are less restrictive than BMA (0-299 for own and spill-in research, 0-107 for own extension). The variances of the proposal distributions are set so that the acceptance rate[28] is close to 25% that is suggested by the literature (Roberts, Gelman and Gilks, 1997).

Baldos, et al. (2018) do not discuss the convergence of estimation; they select the number of iterations without much justification. To make sure that the estimation converges, we check the posterior means of the parameters every 10,000 iterations and stop when the relative change in the posterior means consecutively falls below 1%.

2.7 Postestimation

This section introduces the calculation of elasticities, internal rates of return (IRR) and modified IRR (MIRR) of agricultural R&E, as well as their distributions. Properties of elasticities, IRR, and MIRR are discussed.

[28] The acceptance rate is the percentage that new draws are accepted in M-H sampling.

Calculation of elasticities, IRRs, and MIRRs

Based on coefficient estimates, we calculate the elasticities of MFP to own research, own extension, and spill-in research knowledge stocks (evaluated at the means of the knowledge stocks) and variances[29]:

$$\varepsilon_{\overline{OWN}} = \frac{\partial \ln MFP}{\partial OWN} \cdot OWN\Big|_{\overline{OWN}} = \left[\hat{\beta}_1 + 2\hat{\beta}_4(\overline{OWN} + \overline{EXT}) + \hat{\beta}_6\overline{SPILL}\right] \cdot \overline{OWN} \quad (2.13a)$$

$$\varepsilon_{\overline{EXT}} = \frac{\partial \ln MFP}{\partial EXT} \cdot EXT\Big|_{\overline{EXT}} = \left[\hat{\beta}_1 + 2\hat{\beta}_4(\overline{OWN} + \overline{EXT}) + \hat{\beta}_6\overline{SPILL}\right] \cdot \overline{EXT} \quad (2.13b)$$

$$\begin{aligned}\varepsilon_{\overline{SPILL}} &= \frac{\partial \ln MFP}{\partial SPILL} \cdot SPILL\Big|_{\overline{SPILL}} \\ &= \left[\hat{\beta}_2 + 2\hat{\beta}_5\overline{SPILL} + \hat{\beta}_6(\overline{OWN} + \overline{EXT})\right] \cdot \overline{SPILL}\end{aligned} \quad (2.13c)$$

It is obvious that $\varepsilon_{\overline{OWN}}$ and $\varepsilon_{\overline{EXT}}$ are always of the same sign, so their *p*-values should be identical. Upon calculation of elasticities, for obtaining the IRR of own research, a marginal increase[30] (one thousand dollars) in the expenditure in year $j = 0$ is assumed. This will increase the knowledge stock in year j by w_j $(j = 0,1,\dots,T-1)$. The proportional changes in own research knowledge stock and MFP are $w_j/\overline{OWN}$ and $\varepsilon_{\overline{OWN}} \cdot w_j/\overline{OWN}$. As a result, agricultural output in year j increases by $\varepsilon_{\overline{OWN}} \cdot w_j/\overline{OWN} \cdot \bar{Q}$. $\bar{Q}$ is the mean of total agricultural outputs proxied by real cash receipts. Assuming a discount rate of r, the present value of the increase in outputs (total benefits) is:

$$\sum_{j=0}^{T-1} \varepsilon_{\overline{OWN}} \cdot \frac{w_j}{\overline{OWN}} \cdot \bar{Q} \cdot \frac{1}{(1+r)^{j+1}} = \varepsilon_{\overline{OWN}} \cdot \frac{\bar{Q}}{\overline{OWN}} \sum_{j=0}^{T-1} \frac{w_j}{(1+r)^{j+1}} \quad (2.14)$$

Since the cost is one, IRR is obtained by solving

$$\sum_{j=0}^{T-1} \frac{w_j}{(1+r)^{j+1}} = \frac{1}{\varepsilon_{\overline{OWN}}} \cdot \frac{\overline{OWN}}{\bar{Q}} \quad (2.15)$$

[29] $CLIM$ is unnecessary in Equations 2.13 because its mean is already zero. When calculating the variances of elasticities, we approximate *t*-distribution of the coefficients with normal distribution for simplicity.

[30] Actually, when calculating IRR in this way, the value of the marginal change does not influence the results if the change is small.

The IRR of own extension expenditures is calculated similarly. Calculating the IRR of spill-in research is of little sense in this study because the benefits that Virginia's agriculture receives from other states' research is only a small fraction of the total contributions. As a result, the estimated IRR of spill-in research considerably underestimates its benefits.

The concept of IRR is subject to several flaws, and a modified IRR (MIRR) is suggested to address two of them (Lin, 1976; Hurley, Rao and Pardey, 2014). First, IRR assumes that benefits are reinvested and the rate of return is the same as the original investments. Second, multiple IRRs might exist if net benefits are negative in some year. That could yield ambiguous conclusions about accepting or rejecting an investment project. The second problem is irrelevant because our method of calculating IRR offers either a single solution or no solution as proved below. The first problem could have a profound influence since it is unlikely that the benefits of agricultural research could be invested in research again and receive the same rate of return.

For calculating MIRR, we explicitly determine a finance rate of costs and a reinvestment rate of benefits and calculate the present value of costs and the future value of benefits. Assuming we invest the PV of costs at the beginning and receive the FV of benefits at the end, MIRR is defined as the rate of return to the PV of costs. Given the reinvestment rate r_0, the FV of the benefits of own research is

$$FV = \sum_{j=0}^{T-1} \varepsilon_{\overline{OWN}} \cdot \frac{w_j}{\overline{OWN}} \cdot \bar{Q} \cdot (1 + r_0)^{T-j-1} = \varepsilon_{\overline{OWN}} \cdot \frac{\bar{Q}}{\overline{OWN}} \sum_{j=0}^{T-1} w_j (1 + r_0)^{T-j-1} \qquad (2.16)$$

Since the PV of costs is one, MIRR of own research is

$$MIRR_{OWN} = \left[\varepsilon_{\overline{OWN}} \cdot \frac{\bar{Q}}{\overline{OWN}} \sum_{j=0}^{T-1} w_j (1 + r_0)^{T-j-1} \right]^{\frac{1}{T}} - 1 \qquad (2.17)$$

MIRR of own extension is calculated similarly. Some properties of the IRR and MIRR obtained by the above approach need to be mentioned. First, noticing that the LHS of Equation 2.15 is a monotonically decreasing function of r that approaches zero as $r \to \infty$ and approaches infinity as $r \to -1$, it is easy to prove that IRR increases as elasticity increases and knowledge stock decreases, which is consistent with intuition. Second, a single solution of IRR can be obtained in the range $(-1, \infty)$ when $\varepsilon > 0$, while solution does not exist when $\varepsilon \leq 0$[31]. Negative solutions are possible and imply that, though R&E activities enhance productivity, the gains cannot make up for the costs. Thirdly, like elasticities, IRRs of research and extension are always of the same sign and *p*-value. To prove that, notice that the RHS of Equation 2.15 is identical for own research and extension:

$$\frac{\overline{OWN}}{\varepsilon_{\overline{OWN}} \cdot \bar{Q}} = \frac{\overline{EXT}}{\varepsilon_{\overline{EXT}} \cdot \bar{Q}} = \frac{1}{\left[\hat{\beta}_1 + 2\hat{\beta}_4(\overline{OWN} + \overline{EXT}) + \hat{\beta}_6 \overline{SPILL}\right]\bar{Q}} \equiv \frac{1}{Q^*} \tag{2.18}$$

Therefore, $IRR_{OWN} = IRR_{EXT} = 0$ if $Q^* = 1$. Since LHS of Equation 2.15 is a decreasing function of r, both IRR_{OWN} and IRR_{EXT} are positive if $Q^* > 1$ and negative if $Q^* < 1$. Finally, MIRR=IRR if r_0=IRR; the proof is trivial. Since MIRR is an increasing function of r_0, MIRR<IRR if r_0<IRR. We set $r_0 = 3\%$, which is much smaller than IRRs commonly reported in the literature. Therefore, estimated MIRRs will be less than IRRs, also predicted by Hurley, Rao and Pardey (2014).

Distributions of elasticities, IRRs, and MIRRs

Obtaining the distributions of the postestimation quantities is straightforward for BHM: we calculate the quantities for each draw of parameters and receive a simulated sample of the quantities. The steps to obtain the distributions of the quantities for BMA are shown in the right part of Figure 2.1. For BMA, the uncertainty of a Δ comes from two sources: the uncertainty of models and the uncertainty of coefficients given a model. Therefore, at first, models are repeatedly drawn from the set of valid models using the posterior model probabilities as sampling weights. For each model drawn, a set of simulated coefficients is drawn from its distribution[32], and elasticities, IRRs, and MIRRs are calculated. When $\varepsilon \leq 0$, IRR and MIRR do not exist, and we let them to be -1, which are their limits as $\varepsilon \to 0$, to "hold the seat"[33]. Similar to estimation, convergence is checked after every 10,000 additional draws. Convergence is attained when the changes of the 10%, 25%, 50%, 75%, and 90% quantiles between two checkpoints fall below $1 \times$

[31] We consider only real solutions. Complex solutions are always available.

[32] Theoretically, OLS coefficient estimates are *t*-distributed. For simplicity, we draw simulated coefficients from a multivariate normal distribution.

[33] Actually, when $\varepsilon < 0$, we could calculate MIRR if lag length T is odd, and the resulted MIRR<-1. However, whether T is even or odd should not have an important influence on the conclusions, and we let MIRR$=-1$ when $\varepsilon < 0$ even that T is odd.

10^{-5}. For testing the null hypothesis that Δ>0, we calculate the empirical *p*-value, which is the proportion of the drawn Δ's that are negative.

2.8 Data

This section introduces the data and methods used to compute agricultural the multi-factor productivity (MFP) index. Trends in MFP and agricultural R&E expenditures are discussed and compared for Virginia and its neighbor states.

Multi-factor productivity index (MFP)

Two time series of state agricultural productivity indices are widely used: the total-factor productivity index (TFP) by USDA/ERS and the multi-factor productivity index (MFP) by International Science & Technology Practice & Policy (InSTePP) (Alston, et al., 2010). We employ the latter as the dependent variable because it covers a longer period: 1949-2007 vs. 1960-2004. Version 5 of InSTePP US Production Accounts include national and state-specific estimates of aggregated agricultural input, output, and MFP indices spanning 1949-2007. We follow the same procedure used by InSTePP and extend the series to 2016 for Virginia using data from various sources (e.g., Census of Agriculture, USDA/NASS Quick Stats, Virginia Agricultural Statistics Bulletin). Virginia's output quantity index includes 38 items, accounting for ≈90% of the state's agricultural output by market value during 2000-2016. The input index includes 28 agricultural inputs belonging to four major groups (land, labor, capital, and materials). The MFP is calculated as the ratio of the output to the input index.

Virginia's agricultural productivity grew at an average annual rate of 1.3% during 1949-2016, making the level of productivity in 2016 2.4 times that in 1949 (Figure 2.2). Productivity growth has effectively stagnated since 1990. During 1949-90, the average annual growth rate of Virginia's MFP is 2.1%, which is almost the same as the national rate, but lower than North Carolina (3.0%) and Maryland (2.4%). Productivity growth in Virginia turned negative (-0.9%) from 1991 to 2007, while the national rate (1.2%) and rates of NC (0.7%) and MD (0.9%) declined compared to early periods, but remained positive. Since 2008, the growth in Virginia experienced a recovery, and the average growth rate turned positive (1.4%). A key characteristic of the trend of inputs in Virginia (and other states) is the increase in materials and the decrease in other inputs, especially labor (Figure 2.3). For outputs, greenhouse and nursery products grow rapidly, though their share in Virginia's agriculture is still limited. The output of livestock increases slowly and steadily, while the outputs of field crops, vegetables, and fruits and nuts are volatile.

Agricultural R&E expenditures

The data on public expenditures on agricultural research in Virginia and relevant states come from various USDA reports: Report on the State Agricultural Experiment Stations (1925-59), Funds for Research at State Agricultural Experiment Stations (1960-73), Inventory of Agricultural Research (1974-97), and summary reports of Current Research Information System (CRIS, 1998-2016). Agricultural research expenditures include funds to SAES and 1890 land-grant institutions, which are major practitioners of state agricultural research, and exclude expenditures on forestry and wildlife, which do not directly contribute to agriculture. The data of expenditures on extension services in Virginia come from Norton and Paczkowski (1993) and Annual Reports of Cooperative Extension/Agricultural Experiment Station Division by Virginia Tech. To obtain the real values of R&E expenditures, we assume that 30% of the expenditures are spent on purchasing goods and services and the rest is spent on labor. A deflator is created by using the price index of government purchase of intermediate goods and service from the Bureau of Economic Analysis (BEA) and CPI of urban residents as price indices for the two components.

Real public expenditures on agricultural research in Virginia grew at an average annual rate of 4.4% during 1925-2016, which is higher than the national rate (3.2%) and neighboring states except for Tennessee (4.5%) (Figure 2.4). Similar to productivity, the growth of research expenditures decelerated in recent decades for Virginia, its neighbors, and the U.S. as a whole. In Virginia, the growth rate was 7.4% in 1925-70, dropping dramatically to 1.7% in 1970-2016. Although Virginia's expenditure on extension services was higher than research initially, its growth is slower: the average annual growth rate in 1941-2016 is 1.8%. The growth rate of extension expenditures also declines with time: 5.4% in 1941-1970, 1.1% in 1971-1990 and negative in 1991-2016.

2.9 Results

The model space of BMA analysis includes over 2,168 million combinations of covariates and the same number of models. About 112 million models are drawn until convergence.

Approximately 13 million draws pass the SA tests, and 44,725 are unique models[34]. Not every SA model is valid because some give negative elasticities, implying an adverse influence of R&E activities on productivity. We consider two groups of models that do not conflict with prior knowledge. Group A excludes models that give significantly negative[35] ε_{OWN}, ε_{EXT}, and ε_{SPILL}. Group B, a subset of Group A, keeps models in Group A that give significantly positive ε_{OWN} and ε_{EXT}. Groups A and B include 2,031 and 365 models, respectively. The estimation of BHM generates >38 million draws until convergence and takes a much longer time than BMA. The means and medians of the parameters as the number of draws increases are plotted in Appendix Figure 2.3.

Characteristics of models selected by BMA

Although the pre-determined BMA model space contains a large number of models, models that pass the SA tests and coincide with prior knowledge are in a restricted sub-space (Table 2.3). All models in Groups A and B and almost all the SA models include $SPILL$ based on neighboring states, a linear time trend and $SPILL^2$ but not include the climate variable. The majority of selected models include $(OWN + EXT)^2$ and $(OWN + EXT) \times SPILL$. Therefore, the two criteria of model selection identify a particular function form, which is adopted for BHM (Equation 2.11). The exclusion of climate is surprising at first glance but understandable given the structure of Virginia's agriculture. By cash receipt, livestock accounts for about 61% of Virginia's agricultural output during 1949-2016, and the share increases with time (66% in 2000-2016). Livestock production is less dependent on weather compared with crops and supplies an explanation to the exclusion of weather from models.

Estimated lag structures

In addition to functional form, model selection criteria in BMA analysis also select particular lags. Table 2.4 presents the posterior inclusion probabilities[36] of lag structures by b parameter and peak for all SA models and models in Groups A and B. The main difference between

[34] We run our code on a computer with Intel i7-4790 CPU and 16GB memory, and employ the parallel computing technology (six threads) to accelerate estimation. The estimation takes more than a whole day to converge. That means estimating all the models in the model space will take more than 20 days.
[35] To be consistent with the statistical adequacy tests, we determine the significance at the level of 20%.
[36] The posterior inclusion probability of a covariate or lag structure is the percentage of models that include the covariate or lag structure. For instance, 37.6% of the 44,725 statistically adequate models include own research knowledge stocks based on lag structures with the b parameter in 0.2-0.8.

all SA models and Group A/B is the peak of lags: the posterior inclusion probabilities by peaks are quite concentrated by Group A/B, particularly for own and spill-in research. For own research, 67.8% of Group A models and 88.2% of Group B use lags peaking at $j = 4$. For own extension, 74.9% of Group A and 61.6% of Group B use lags peaking at $j = 5$ or $j = 6$. For spill-in research, all models in both groups use lag structures peaking at year 15-18. The b parameter is important for spill-in research: it tends to be small for all SA models (flat lag structure) and large for Group A/B (steep lag structure).

Since the peaks are less restricted in BHM estimation, the posterior distributions of the peaks have very long tails on the right tail, while the distribution of b parameter is relatively flat (Figure 2.5). The mean, median, and standard error of the posterior distributions of b parameter and peak are shown in Table 2.5. Summary statistics of BMA are obtained similarly as that of elasticities and rates of return. The parameters are possibly dependent. For instance, Figure 2.6 includes the pairwise scatter plots of the lag parameters drawn in BHM estimation[37], showing obvious linear relationships between a_{EXT} and b_{EXT} and between a_{SPILL} and b_{SPILL}. Therefore, the marginal mean and median are probably improper measures of the central tendency, and we resort to the concept of spatial median. The spatial median, or multivariate median, is the median of multi-dimensional data that minimizes the average Euclidean distance between the point of spatial median and the data points in the sample (Vardi and Zhang, 2000).

The median and spatial median of lag peaks are identical by BMA: 4 for own research, 5 for own extension, and 16 (Group A) or 17 (Group B) for spill-in research. By BHM, the median peaks of own research and extension are almost the same (6.3), while by spatial median own research exhibits its maximum effects slightly later than extension (7.6 vs. 6.9). Therefore, BHM offers later peaks of own research and extension than BMA. By contrast, BHM predicts a much earlier peak (11.3) of spill-in research than BMA.

Figure 2.7 shows the shape of the lag structures based on the spatial median of the lag parameters, referred to as *typical lags* below. To present the uncertainty of the shapes, 95% "confidence intervals" of lag weights are obtained for each j, which is the (2.5%, 97.5%) quantiles of lag weights. The lag structures estimated by BHM are more uncertain than those by BMA, because of BHM's less restricted parameter space. For own research, the typical lags by the two

[37] It is impossible to plot all the 38 million draws in the graph, so we randomly draw 10,000 out of the 38 million draws and plot them. We have tried several times and the patterns are stable.

groups of BMA are almost identical, while that by BHM is flatter and takes more time to reach the peak and diminish to zero. For own extension, the typical lag weight does not decline to zero at $j = 8$, suggesting the need for a longer lag to fully capture benefits of extension. The key difference between BMA and BHM happens for spill-in research. The typical lag structure of spill-in research by BMA has a gestation period of about eight years, peaks at about $j = 16$, and the lag weight diminishes to almost zero at $j = 24$. By contrast, the typical spill-in research lag by BHM has a short gestation of about four years, peaks at about $j = 10$ and diminishes to zero at $j = 20$.

Elasticities and rates of return

Summary statistics of elasticities and rates of return are presented in Tables 2.6 (elasticities), 2.7 (IRR), and 2.8 (MIRR); their distributions (density plots) are shown in Figures 2.8 (elasticities) and 2.9 (IRR and MIRR). The distributions of the elasticities given by BHM is more diffuse (flat) than BMA and slightly to the right. The flatter distributions by BHM are also due to fewer restrictions on parameters.

The mean and median elasticities by BHM are larger than BMA, while the two groups of BMA are quite similar. For IRR and MIRR, the asymmetry of the distributions, the spike at -100%, and the long right tail imply that discussing the mean is not sensible (Figure 2.9). The median IRRs of own research and extension are 12.9% and 11.7% by Group A of BMA, 38.0% and 31.2% by Group B of BMA, and 19.8% and 43.1% by BHM. MIRRs are smaller than IRRs as predicted: 5.2% and 8.9% by Group A of BMA, 9.3% and 21.2% by Group B of BMA, and 11.9% and 31.1% by BHM. The spatial medians of the rates of return by Group A of BMA are close to zero or negative. By Group B of BMA, the spatial medians (IRR: 35.8% and 28.8%, MIRR: 8.0% and 20.5%) are slightly lower than the marginal medians. For BHM, the spatial medians of MIRRs and the IRR of own extension are close to the marginal medians, while the spatial median of own research IRR (26.4%) is larger than the marginal median. When marginal and spatial medians are different, the spatial median is preferable because it is a more robust estimator of the central tendency.

The empirical p-value of IRR is 0.434 for Group A of BMA, 0.194 for Group B of BMA, and 0.304 for BHM[38]. For MIRR, the empirical p-values of own research and extension are 0.406 and 0.429 by Group A of BMA, 0.171 and 0.190 by Group B of BMA, and 0.297 and 0.300 by

[38] We have proved that, for IRR, the empirical p-value of research and extension should be the same.

BHM. Therefore, the rates of return given by Group B of BMA are more statistically "significant" than Group A and BHM. Empirical studies usually adopt 1% or 5% as the threshold to determine the significance of statistics. The *p*-values in this study are quite large and will not be taken as significant in most studies. However, we should remember that the majority of empirical studies on returns to R&E rely on a single ad hoc lag structure to obtain the confidence intervals of estimates, so the uncertainty of the lag structures is ignored. By contrast, the results of this study consider the uncertainty of lag structures and should be more uncertain.

Since state-level evaluations, particularly for Virginia, are rare[39], only a few studies can be used for comparison. Norton and Paczkowski (1993) use a 12-year polynomial lag for research and 9-year polynomial lag for extension and cover the period 1949-89; their estimated IRRs are 58% and 37% for research and extension respectively. Alston, et al. (2010) do not distinguish extension from research and use a 50-year Gamma lag; the IRR of R&E combined is 26.1% by a linear model and 16.7% by a log model in 1949-2002. Plastina and Fulginiti (2012) use an inverted-V lag of 31 years and cover the period 1949-91. They estimate a translog cost function and ignore extension, and the estimated IRR of research is 17.9%. These studies use different lags and evaluation methods and cover different periods, which makes the results not comparable. However, we question the reliability of the latter two multi-state studies because they apply the same lag structure to all states. In particular, the lag structure used by Plastina and Fulginiti (2012) is from a nonparametric analysis based on national data by Chavas and Cox (1992). We do not believe it is valid to apply the lag structure estimated from national data to state-level analysis.

Growth accounting

Growth accounting decomposes the productivity growth by its sources. The following expression for the change in $\ln MFP$ between 1949 and 2016 can be derived:

$$\begin{aligned}\Delta \ln MFP \equiv{}& \ln MFP_{2016} - \ln MFP_{1949} \\ ={}& [(\beta_1 + 2\beta_4\overline{EXT} + \beta_6\overline{SPILL})\Delta OWN + \beta_4\Delta OWN^2] + \\ & [(\beta_1 + 2\beta_4\overline{OWN} + \beta_6\overline{SPILL})\Delta EXT + \beta_4\Delta EXT^2] + \\ & [(\beta_2 + \beta_6\overline{OWN} + \beta_6\overline{EXT})\Delta SPILL + \beta_5\Delta SPILL^2] + 67\beta_9\end{aligned} \tag{2.19}$$

39 Although several studies use state-level data, they do not report the rates of return for each state.

In Equation 2.19, $\Delta OWN \equiv OWN_{2016} - OWN_{1949}$, $\Delta OWN^2 \equiv OWN_{2016}^2 - OWN_{1949}^2$, $\overline{OWN} \equiv (OWN_{1949} + OWN_{2016})/2$; other variables are defined similarly. Equation 2.19 decomposes $\Delta \ln MFP$ into four components: the growth that resulted from OWN, EXT, $SPILL$, and time trend. Since time trend represents unobserved technology, $SPILL$ and time trend are combined into a single component reflecting total spillover effects.

The distributions of the fractions of MFP growth from the three sources are obtained similarly as rates of return. Table 2.9 and Figure 2.10 present summary statistics and density plots. The fractions can be negative or over 100%. However, the means and medians[40] of the fractions are always in the (0, 100%) range and offer a good summary of the results. By the medians, the fractions of Virginia's agricultural productivity growth in 1949-2016 that can be attributed to its investments in research and extension are 6.7% and 3.1% by Group A, 14.0% and 6.6% by Group B, and 12.7% and 6.3% by BHM. Group B of BMA and BHM give similar results: about 20% of Virginia's agricultural productivity growth comes from its R&E activities.

Only one study can be used to compare the results. By Alston, et al. (2010)[41], the linear model implies that 23.5% of Virginia's agricultural productivity growth comes from its R&E expenditures, which is close to the results of this study. However, the logarithmic model preferred by authors offers a much larger estimate (63.2%). Virginia's agriculture benefit not only from its agricultural R&E but from those in other states and countries. Agricultural productivity benefits not only from the research in agriculture but also from related areas like biological and environmental sciences. Therefore, Virginia's agricultural R&E activities are only a small fraction of the total R&E efforts that contribute to Virginia's agriculture. This fact implies that 63.2% most probably exaggerates the contribution, and 20% is a more reasonable estimate.

Simulations

Based on the estimation results, the effects of own research/extension investments are investigated in four simulated scenarios[42]: 1) one million dollars extra expenditure on own research or extension in 1992; 2) one million dollars extra expenditure on own research or extension in 1983; 3) maintain the growth rate of own research or extension expenditure in 1992-2016 as that

[40] Since the fractions of three sources sum up to one, the spatial medians are always identical to marginal medians.
[41] The study use a different approach for growth accounting: they do growth accounting for each pair of successive years and take the average of the percentages.
[42] The years are selected because they are turning points in the trend of MFP or expenditures.

in 1971-91 (2.7% for research, 0.7% for extension); and 4) maintain the growth rate of own research or extension expenditure in 1992-2016 as that in 1949-91 (4.9% for research, 3.2% for extension). Since the growth of agricultural productivity and R&E expenditures decelerate together in recent decades, Scenarios 3 and 4 are particularly interesting.

The benefit/cost ratios in Scenarios 1 and 2 are calculated. For Scenarios 3 and 4, since the benefits last beyond 2016, the total benefits cannot be calculated, and the focus is on the change in the MFP in 2016. In general, the following relationship holds for the benefit/cost ratios in Scenarios 1 and 2 and the percentage change in 2016 MFP in Scenarios 3 and 4: Group A of BMA < Group B of BMA < BHM, and the mean, median, and spatial median are similar (Table 2.10). The benefit/cost ratio in Scenario 1 is slightly lower than that of Scenario 2. Since the assumed extra investments are later in Scenario 1 than in 2, the results confirm that the returns to R&E declines with time. For Scenarios 3 and 4, if the expenditure on own extension maintains the historical growth rate, the change in 2016 MFP is larger than if own research maintains its growth rate, which seems odd because BMA offers a higher IRR of research. These results arise because the investments in own extension show a more serious recession than research in recent decades. Maintaining the historical growth rate means larger extra investments in extension than in research.

Relationship between MIRR and reinvestment rate

By the results of BMA, the relative magnitudes of the rates of return of research and extension are different by IRR and MIRR: median $IRR_{OWN} > IRR_{EXT}$ but median $MIRR_{OWN} < MIRR_{EXT}$. To find out the reason, we estimate models using the typical lag structures in Figure 2.7 and calculate MIRR with various reinvestment rates. The relationships between MIRR and reinvestment rate are shown in Figure 2.12.

The relationship between MIRR and the reinvestment rate is almost linear, and as the reinvestment rate decreases, MIRR of research drops more quickly than that of extension. Therefore, the two curves cross and the relationship between $MIRR_{OWN}$ and $MIRR_{EXT}$ flips at some point (≈11% for Group A of BMA, ≈34% for Group B of BMA, ≈70% for BHM). This flip could be proved informally. If $r_0 = 0$ in Equation 2.17, $MIRR = Q^{*1/T} - 1$. Since $T_{OWN} > T_{EXT}$, $MIRR_{OWN} < MIRR_{EXT}$ if $r_0 = 0$. We have proved that $MIRR = IRR$ if $r_0 = IRR$. Therefore, as $IRR_{OWN} > IRR_{EXT}$ by BMA, if we evaluate MIRR at $r_0 = IRR$, $MIRR_{OWN} > MIRR_{EXT}$. There must be a r_0 between zero and IRR where the relationship between $MIRR_{OWN}$

and $MIRR_{EXT}$ flips. For BHM, since the median IRRs are much lower than the flipping point, we observe the same relationship between IRR_{OWN} and IRR_{EXT} and between $MIRR_{OWN}$ and $MIRR_{EXT}$. This result can be understood in the context of compounding: reducing a rate of return will lead to a larger impact on a longer investment program. This is consistent with the prediction by Hurley, Rao and Pardey (2014) that more profitable investments tend to generate larger difference between IRR and MIRR.

BMA without SA tests

To understand the role of SA tests in BMA analysis, we skip the tests and repeat the estimation procedure[43], and the results are fundamentally different. First, the posterior inclusion probabilities of covariates are diffuse. Even if we filter the models with prior knowledge, each lag structure is included almost with the same probability. Therefore, without SA tests, prior knowledge does not identify a particular functional form or lag structures. Second, the ranges of estimated rates of return grow substantially. For instance, the 99% confidence intervals of own research and extension IRRs by Group B become (-100%, 153.1%) and (-100%, 673.6%), compared with only (-100%, 101.0%) and (-100%, 91.1%) when tests are not skipped. Thirdly, the medians of estimated rates of return also increase by a large extent. For Group B, median IRRs of R&E increase from 38.0% and 31.2% to 73.3% and 86.2%, and median MIRRs increase from 9.3% and 21.2% to 15.2% and 40.5%. Finally, while own research IRR is larger than own extension IRR with SA tests, the reverse is true if tests are skipped. Tests of SA tend to rule out models giving large rates of return, particularly those giving large rates of return of extension.

Testing assumptions of BHM

Like most empirical studies, studies using Bayesian techniques usually fail to test statistical assumptions underlying the models. The GS method and M-H algorithm provide not only a simulated sample of parameters but also a simulated sample of residuals. Using simulated residuals, testing the assumptions is straightforward. A problem with testing the simulated residuals is that the huge number of draws makes any difference significant, so null hypotheses are easily rejected. For the first five observations (1949-53), we calculate the residuals by each set of drawn

[43] Posterior inclusion probabilities and summary statistics of elasticities and rates of return are presented in Appendix Tables 2.1, 2.2, and 2.3.

parameters and test the normality and whether the mean is zero. The null hypotheses are always rejected as expected (Table 2.11). However, if we do not look at the *p*-values, the mean of the residuals is quite close to zero for each year, and the skewness and kurtosis are very close to those of a normal distribution. It is the huge sample size making the small difference statistically significant.

Alternatively, we can test the residuals analogously to the test of an OLS regression: calculating residuals based on a particular point in the parameter space and testing the realized residuals. When testing an OLS regression, we calculate the residuals based on the mean of the parameters. Given a CLRM with NIID error terms, the residuals are normally distributed, so mean and median are identical and sound measures of central tendency. Since the drawn parameters of BHM do not follow any known distribution, we calculate residuals based on three representative points (marginal mean, marginal median, and spatial median), do the SA tests, and calculate elasticities, rates of return, and growth accounting fractions (Table 2.12). The normality assumption cannot be rejected at all three points. The spatial median of parameters almost passes the joint test of conditional variance if using 20% as the threshold. Therefore, by the spatial median, which is a more robust measure of central tendency, BHM estimation results are sound on normality tests and conditional variance tests. This result is acceptable because the key assumption of BHM is the likelihood function that results from the assumptions on the error terms. Failing to reject the null hypotheses of the conditional mean test implies there is unexplained information in the residuals.

Choice of the final result

We prefer Group B of BMA to Group A because Group B models predict better. How well models predict is another way of evaluating the roles of SA tests and prior knowledge in the model selection strategy. For each model, the observations of the first 63 years are used for estimation and the remainder is used for prediction. Table 2.13 shows the percentages of models whose 95% CI of prediction contain the actual value for each year in 2012-16. The last row shows the percentage of models whose prediction CIs in all five years contain the actual values. Models that pass SA tests predict much better: 72.6% offer prediction CIs containing the actual value for all five years, compared with only 33.6% when tests are skipped. Group A is a little worse on prediction than all SA models (64.9%), while Group B is better than Group A and all SA models (90.1%).

The spatial medians of own research and extension IRRs given by Group B of BMA are 35.8% and 28.8%, while those by BHM are 26.4% and 42.4%. BHM differs from BMA in that BHM imposes fewer restrictions on the lag structures. The peaks of lag structures influence rate-of-return estimates substantially. The distribution of the own research peak concentrates at early years by both BMA and BHM. However, while the majority of models in Groups A and B of BMA use own research lags peaking at $j = 4$, 46.0% of BHM draws are with own research peak>7 (Table 2.14), and the IRR of own research tends to be smaller with later peaks. The peak of own extension lags does not have a large effect on the IRR of research, but the effect on the IRR of extension is large as expected. The peaks of extension lags selected by BMA are from year 5 to 8, while 26.0% of the BHM draws gives peak<=4 with an extremely large median IRR of 86.6%. For spill-in research, SA tests in BMA exclude almost all lags with peak <9 or >18. By BHM, peaks of 44.2% of the draws are <=9, which gives a median own extension IRR of 56.3%. Therefore, a considerable share of BHM draws are out of the model subspaces of Groups A and B, and these draws in general give smaller research IRR and larger extension IRR. As a result, we observe smaller research IRR estimates and larger extension IRR estimates by BHM compared with BMA.

We have more confidence in the results given by BHM. First, we want to impose less structure on the model and let the data speak for econometric modeling. This principle favors BHM because the lag structures are less restricted. Second, the typical lags of own research and extension are similar by BMA and BHM, and the typical spill-in research lag given by BHM is more reasonable than that by BMA. The spill-in research knowledge stock in this study includes only research of neighbor states. Therefore, the spill-in research is expected to take effect with a slightly longer lag than own research. The typical lags by BHM conform with this expectation while the spill-in research lag by BMA reveals a gestation period that is too long to be true. Thirdly, although typical lags of own research and extension look similar by BMA and BHM, a subtle difference makes BHM preferable: BHM predicts an earlier peak of extension than research, while BMA gives the opposite. Extension is expected to take effect more rapidly than research because it is more locally oriented. Finally, it is expected $IRR_{OWN} > IRR_{EXT}$, so the results by BMA seems more reasonable. However, we should keep in mind that the spillover effects of Virginia's agricultural research have been neglected. Therefore, the estimated own research IRR underestimates the benefits of Virginia's research investments.

2.10 Conclusion and methodology revisited

This study estimates rates of return of Virginia's public expenditures on agricultural R&E from 1949 to 2016. We compare econometric modeling strategies in previous studies and conclude that BMA and BHM are promising means of solving the fundamental problem of model uncertainty, although each method has its pros and cons. BMA depends on a pre-determined set of lag structures but allows various functional forms. BHM imposes fewer restrictions on the lag structures but requires prior specification of a functional form. When the BMA model space is filtered using SA tests and prior knowledge, a particular functional form emerges and can be applied to BHM, which provides the incentive of combining the two methods.

The characteristics of the estimated research lags are consistent with expectation: own research takes effect promptly, while research in other states needs more time to influence productivity. The length of own research lag is about 15 years by BMA and 18 years by BHM, and the lag reaches its peak at $j = 4$ by BMA and $j = 7.6$ by BHM. The estimated own extension lag needs about 15 years to diminish to zero. Therefore, lags longer than nine years are needed to capture the effects of extension fully. Although the typical lags of own research and extension are similar by BMA and BHM, the typical spill-in research lag by BHM has a shorter gestation and peaks earlier than that given by BMA.

The truncated own extension lag calls for a longer lag, but using a longer lag means dropping some observations from the beginning of the time series. We attempted to extend the own extension lag to $T = 12$ and repeated the BMA analysis. Although only three observations are dropped, the results change greatly, and the main reason is a different set of models pass the SA tests. For instance, while almost all the models include the linear time trend when $T = 9$, about half of the models passing SA tests included a quadratic time trend. By Figures 2.2 and 2.4, the first and second halves of the time series of MFP and R&E expenditures are fundamentally different: the first half shows a trend of steady growth, while the second half is flat and volatile. Randomly dropping a few observations may not greatly alter the results of a regression using cross-sectional data, but dropping observations from the beginning of a time series with a structural change could change the result dramatically. Therefore, detecting the influential observations or "outliers" and estimating models when outliers present is the topic of a large number of time-series studies (e.g., Chang, Tiao and Chen, 1988; Ljung, 1993). The methodology of this study judges the results from several aspects: the models are expected to be statistically sound, the shapes of the estimated lags are expected to be in line with expectation, the estimated rates of return and growth accounting results are consistent with prior knowledge, etc. It is very difficult to make the results sound on every aspect. We prefer the truncated own extension lag because the current results are the best in terms of fulfilling the criteria above, accounting for trade-offs.

BHM relies on fewer restrictions on lag structures and offers estimated lag structures that are more reasonable. By BHM, the spatial median IRRs are 26.4% and 42.4% for own research and extension, respectively. By growth accounting, 19.0% of Virginia's agricultural productivity growth in 1949-2016 comes from its R&E activities, and the contribution of the research is about twice of extension. One extra million dollar expenditure on research in 1992 will bring a benefit of $4.5 million, and the same expenditure in 1983 will bring $5.4 million. If the extra expenditure is spent on extension, it will bring a benefit of $6.1 million or $6.3 million if the expenditure happens in 1992 or 1983. MFP in 2016 will be 3.2% and 8.5% higher if own research expenditure in 1992-2016 maintains the growth rate in 1971-91 and 1949-91. If own extension expenditure maintains the historical growth rate, MFP in 2016 will be 6.7% and 17.5% higher.

The key principle guiding the modeling strategy is the desire to avoid arbitrariness to the degree possible, and the principle leads to choosing BHM over BMA. However, BHM is not superior to BMA in every dimension. First, BMA is much easier to implement and less demanding of computation power. The M-H algorithm used in BHM estimation draws a set of parameters based on the previous draw, so parallel computing technology is not applicable. Besides, the knowledge stocks are taken as parameters instead of data. These issues make BHM converge more slowly than BMA. The huge number of draws generated by BHM also slow the calculation of elasticities, rates of return, and other post-estimation steps.

Second, BMA allows various functional forms, while BHM requires a functional form to be specified prior to the start of estimation. In this study, the SA tests select a particular functional form from the BMA model space, which is not guaranteed in every study. Otherwise, we still need to determine the functional form for BHM.

While model averaging is well-known in the literature, we use specific criteria to limit model space. The underlying assumptions of a model are the foundation of statistical inferences. The concepts of standard errors of coefficients, confidence intervals, and significance are all intimately dependent on a specific set of assumptions. Compared with the SA tests, filtering the model space with prior knowledge seems more problematic because it selects models by results.

We have to admit that economists must use their prior knowledge to build models: using economic theory to specify a model, choosing relevant covariates, etc. However, there is something behind the scene that the practitioners of econometrics often fail to reveal. Practitioners often re-specify the model until they feel the results are consistent with expectations and could pass reviewers' scrutiny. They report their final model in the paper, and sometimes they compare models as means of conducting robustness checks. The majority of the models that fail to pass the process of "try and try again" (Mittelhammer, 1993) are ignored. By contrast, we have presented the pool of models and the criteria of model selection, so our methodology is not worse than the majority of empirical studies at the minimum. If models need to be re-specified, it is preferable to show all models rather than only report the final model.

We believe this approach is superior to the alternatives. First, although we use prior expectation, we use it parsimoniously. We are sure agricultural research should have a non-negative effect on productivity but are not sure about the magnitude. Therefore, models are selected using the sign of the marginal effects, and any magnitude is allowed when the sign is correct. Second, we present results based on different prior expectations (Groups A and B), allowing the reader to choose base on their expectations. For some readers, Group B is probably based on a belief that this too strong to accept, while the belief that Group A emerges from is more acceptable. Therefore, they may prefer the results given by Group A. To sum up, the advantage of model averaging over the "try and try again" approach is its transparency that allows the readers to evaluate the validity of the study to a larger extent.

References

Acquaye, A.K.A., J.M. Alston, and P.G. Pardey. 2003. "Post-War Productivity Patterns in U.S. Agriculture: Influences of Aggregation Procedures in a State-Level Analysis." *American Journal of Agricultural Economics* 85:59-80.

Alston, J.M., M.A. Andersen, J.S. James, and P.G. Pardey. 2010. *Persistence Pays: U.S. Agricultural Productivity Growth and the Benefits from Public R&D Spending*. New York: Springer.

Alston, J.M., C. Chan-Kang, M.C. Marra, P.G. Pardey, and T.J. Wyatt. 2000. "A Meta-Analysis of Rates of Return to Agricultural R&D: Ex Pede Herculem?" IFPRI Research Reports. International Food Policy Research Institute.

Alston, J.M., P.G. Pardey, and H.O. Carter. 1994. "Valuing UC Agricultural Research and Extension." Publication No. VR-1. Agricultural Issues Center, University of California.

Andersen, M.A. 2015. "Public Investment in U.S. Agricultural R&D and the Economic Benefits." *Food Policy* 51:38-43.

Andersen, M.A., and W. Song. 2013. "The Economic Impact of Public Agricultural Research and Development in the United States." *Agricultural Economics* 44:287-295.

Balcombe, K., and G. Rapsomanikis. 2010. "An Analysis of the Impact of Research and Development on Productivity using Bayesian Model Averaging with a Reversible Jump Algorithm." *American Journal of Agricultural Economics* 92:985-998.

Baldos, U.L.C., F.G. Viens, T.W. Hertel, and K.O. Fuglie. 2018. "R&D Spending, Knowledge Capital, and Agricultural Productivity Growth: A Bayesian Approach." *American Journal of Agricultural Economics* 101:291-310.

Ball, V.E., F.M. Gollop, A. Kelly-Hawke, and G.P. Swinand. 1999. "Patterns of State Productivity Growth in the U.S. Farm Sector: Linking State and Aggregate Models." *American Journal of Agricultural Economics* 81:164-179.

Bobb, J.F., F. Dominici, and R.D. Peng. 2011. "A Bayesian Model Averaging Approach for Estimating the Relative Risk of Mortality Associated with Heat Waves in 105 US Cities." *Biometrics* 67:1605-1616.

Brock, W.A., S.N. Durlauf, and K.D. West. 2003. "Policy Evaluation in Uncertain Economic Environments." NBER Working Paper Series. National Bureau of Economic Research.

Bryant, H.L., and G.C. Davis. 2008. "Revisiting Aggregate US Meat Demand with a Bayesian Averaging of Classical Estimates Approach: Do We Need a More General Theory?" *American Journal of Agricultural Economics* 90:103-116.

Casella, G., and E.I. George. 1992. "Explaining the Gibbs sampler." *The American Statistician* 46:167-174.

Chang, I., G.C. Tiao, and C. Chen. 1988. "Estimation of Time Series Parameters in the Presence of Outliers." *Technometrics* 30:193-204.

Chatfield, C. 1995. "Model Uncertainty, Data Mining and Statistical Inference." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 158:419-466.

Chavas, J.-P., and T.L. Cox. 1992. "A Nonparametric Analysis of the Influence of Research on Agricultural Productivity." *American Journal of Agricultural Economics* 74:583-591.

Chib, S., and E. Greenberg. 1995. "Understanding the Metropolis-Hastings Algorithm." *The American Statistician* 49:327-335.

Chipman, H. 1996. "Bayesian Variable Selection with Related Predictors." *Canadian Journal of Statistics* 24:17-36.

Clemen, R.T. 1989. "Combining Forecasts: A Review and Annotated Bibliography." *International Journal of Forecasting* 5:559-583.

D'agostino, R.B., A. Belanger, and R.B. D'Agostino Jr. 1990. "A Suggestion for Using Powerful and Informative Tests of Normality." *The American Statistician* 44:316-321.

Evenson, R. 1967. "The Contribution of Agricultural Research to Production." *Journal of Farm Economics* 49:1415-1425.

Fornash, L. 2011. *Analysis of the Virginia Cooperative Extension Service (VCE) Structure, Funding Trends, and Research*. Richmond, VA: Office of the Governor.

Geisel, M.S. 1973. "Bayesian Comparisons of Simple Macroeconomic Models." *Journal of Money, Credit and Banking* 5:751-772.

Geisser, S. 1965. "A Bayes Approach for Combining Correlated Estimates." *Journal of the American Statistical Association* 60:602-607.

George, E.I. 2010. "Dilution Priors: Compensating for Model Space Redundancy." In *Borrowing Strength: Theory Powering Applications–A Festschrift for Lawrence D. Brown.* Institute of Mathematical Statistics, pp. 158-165.

Greene, W.H. 2012. *Econometric Analysis*. 7th ed. Upper Saddle River, NJ: Prentice Hall.

Griliches, Z. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *The Bell Journal of Economics* 10:92-116.

Hastings, W.K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57:97-109.

Hoff, P.D. 2009. *A First Course in Bayesian Statistical Methods*. New York: Springer.

Huang, Y., F. Dominici, and M.L. Bell. 2005. "Bayesian Hierarchical Distributed Lag Models for Summer Ozone Exposure and Cardi-Respiratory Mortality." *Environmetrics: The Official Journal of the International Environmetrics Society* 16:547-562.

Huffman, W.E., and R.E. Evenson. 2006. "Do Formula or Competitive Grant Funds Have Greater Impacts on State Agricultural Productivity." *American Journal of Agricultural Economics* 88:783-798.

---. 2006. *Science for Agriculture: A Long-Term Perspective*. 2nd ed. Ames, IA: Blackwell Pub.

Hurley, T.M., X. Rao, and P.G. Pardey. 2014. "Re-examining the Reported Rates of Return to Food and Agricultural Research and Development." *American Journal of Agricultural Economics* 96:1492-1504.

Jin, Y., and W.E. Huffman. 2016. "Measuring Public Agricultural Research and Extension and Estimating Their Impacts on Agricultural Productivity: New Insights from U.S. Evidence." *Agricultural Economics* 47:15-31.

Jorgenson, D.W. 1966. "Rational Distributed Lag Functions." *Econometrica* 34:135-149.

Lin, S.A. 1976. "The Modified Internal Rate of Return and Investment Criterion." *The Engineering Economist* 21:237-247.

Ljung, G.M. 1993. "On Outlier Detection in Time Series." *Journal of the Royal Statistical Society: Series B (Methodological)* 55:559-567.

Mayo, D.G., and A. Spanos. 2004. "Methodology in Practice: Statistical Misspecification Testing." *Philosophy of Science* 71:1007-1025.

McGuirk, A.M., P. Driscoll, and J. Alwang. 1993. "Misspecification Testing: A Comprehensive Approach." *American Journal of Agricultural Economics* 75:1044-1055.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. 1953. "Equation of State Calculations by Fast Computing Machines." *The Journal of Chemical Physics* 21:1087-1092.

Mittelhammer, R.C. 1993. "Applications of New Bayesian Techniques to Agricultural Economics: Discussion." *American Journal of Agricultural Economics* 75:1217-1220.

Moral-Benito, E. 2012. "Determinants of Economic Growth: A Bayesian Panel Data Approach." *Review of Economics and Statistics* 94:566-579.

Moral-Benito, E. 2015. "Model Averaging in Economics: An Overview." *Journal of Economic Surveys* 29:46-75.

Norton, G.W., J.D. Coffey, and E.B. Frye. 1984. "Estimating Returns to Agricultural Research, Extension, and Teaching at the State Level." *Southern Journal of Agricultural Economics* 16:121-128.

Norton, G.W., and R. Paczkowski. 1993. "Reaping the Return of Agricultural Research and Education in Virginia." Information Series 93-3. College of Agriculture and Life Sciences, Virginia Tech.

Omernik, J.M., and G.E. Griffith. 2014. "Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework." *Environmental Management* 54:1249-1266.

Pardey, P.G., and B. Craig. 1989. "Causal Relationships between Public Sector Agricultural Research Expenditures and Output." *American Journal of Agricultural Economics* 71:9-19.

Plastina, A., and L. Fulginiti. 2012. "Rates of Return to Public Agricultural Research in 48 US States." *Journal of Productivity Analysis* 37:95-113.

Raftery, A.E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111-163.

Roberts, G.O., A. Gelman, and W.R. Gilks. 1997. "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms." *The Annals of Applied Probability* 7:110-120.

Sala-i-Martin, X., G. Doppelhofer, and R.I. Miller. 2004. "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." *The American Economic Review* 94:813-835.

Schultz, T.W. 1956. "Reflections on Agricultural Production, Output and Supply." *Journal of Farm Economics* 38:748-762.

Shiller, R.J. 1973. "A Distributed Lag Estimator Derived from Smoothness Priors." *Econometrica* 41:775-788.

Thirtle, C., J. Piesse, and D. Schimmelpfennig. 2008. "Modeling the Length and Shape of the R&D Lag: An Application to UK Agricultural Productivity." *Agricultural Economics* 39:73-85.

Vardi, Y., and C.-H. Zhang. 2000. "The Multivariate L1-median and Associated Data Depth." *Proceedings of the National Academy of Sciences* 97:1423-1426.

Welty, L.J., R. Peng, S. Zeger, and F. Dominici. 2009. "Bayesian Distributed Lag Models: Estimating Effects of Particulate Matter Air Pollution on Daily Mortality." *Biometrics* 65:282-291.

Yee, J. 1992. "Assessing Rates of Return to Public and Private Agricultural Research." *Journal of Agricultural Economics Research* 44:35-41.

Zhang, P. 1992. "Inference after Variable Selection in Linear Regression Models." *Biometrika* 79:741-746.

Tables

Table 2.1. Definition of the BMA model space

	Covariate	Possible values	Description
1	OWN	1,2,…,264	264 own research knowledge stocks based on 264 lag structures, obtained by 12 b parameters (0.2, 0.4, …, 2.4) $\times$ 22 peaks (4, 5, …, 25);
2	EXT	1,2,…,72	72 own extension knowledge stocks based on 72 lag structures, obtained by 12 b parameters (0.2, 0.4, …, 2.4) $\times$ 6 peaks (4, 5, …, 9);
3	SPILL	1,2,…,792	792 spill-in research knowledge stocks based on three sets of states $\times$ 264 lag structures; three sets of states are a) five neighbor states (MD, NC, TN, KY, WV), b) 21 states that are similar to VA by agro-ecological conditions (AL, AR, DE, FL, GA, IL, IN, LA, MO, MS, NJ, OH, OK, PA, SC, TX, and the neighbor states), c) all states except VA; lag structures are the same as those used by own research knowledge stock;
4	CLIM	0,1,2	0 for not adding climate variable; 1 for adding precipitation; 2 for adding PDSI;
5	$(OWN+EXT)^2$	0,1	1 if adding $(OWN+EXT)^2$; 0 otherwise;
6	$SPILL^2$	0,1	1 if adding $SPILL^2$; 0 otherwise;
7	(OWN+EXT)×SPILL	0,1	1 if adding (OWN+EXT)×SPILL; 0 otherwise;
8	(OWN+EXT)×CLIM	0,1	1 if adding (OWN+EXT)×CLIM; 0 otherwise;
9	SPILL×CLIM	0,1	1 if adding SPILL×CLIM; 0 otherwise;
10	t, t^2	1,2	1 if adding t; 2 if adding t and t^2.

Table 2.2. Prior and proposal distributions of BHM estimation

	Prior distribution	Proposal distribution
Coefficients of the top model	N($\mathbf{0}$, 100$\mathbf{I}$)*	
Parameters of Gamma distributed lags		
Own research a	TN(7, 100, 0)**	TN(a_{i-1}, 2, 1.8, 61)***
Own research b	TN(1.2, 100, 0)	TN(b_{i-1}, 0.05, 0.2, 2.4)
Own extension a	TN(8.8, 100, 0)	TN(a_{i-1}, 2, 1.2, 22.6)
Own extension b	TN(1.3, 100, 0)	TN(b_{i-1}, 0.05, 0.2, 2.4)
Spill-in research a	TN(16.3, 100, 0)	TN(a_{i-1}, 0.1, 1.8, 61)
Spill-in research b	TN(0.9, 100, 0)	TN(b_{i-1}, 0.0025, 0.2, 2.4)
Variance of error terms (top and bottom models)	IG(0.5, 0.5)	
R&E knowledge stocks		TN($\boldsymbol{S}_{i-1}$, 0.08$\mathbf{I}$, 0)

* $\mathbf{I}$ is an identity matrix.
** TN(μ, σ^2, 0) is a truncated normal distribution with lower bound of zero.
*** TN(μ, σ^2, lwr, upr) is a truncated normal distribution with lower bound lwr and upper bound upr. x_{i-1} is the parameter drawn in the previous iteration.

Table 2.3. Posterior inclusion probabilities of covariates, BMA (%)

	All	Group A	Group B
States for calculating spill-in research knowledge stock			
Neighboring states	99.88	100.00	100.00
Similar by agro-ecological conditions	0.06		
All states except VA	0.06		
Climate variable			
Not included	97.92	100.00	100.00
Precipitation	0.27		
PDSI	1.81		
Time trend			
Linear (only t)	94.15	100.00	100.00
Quadratic (t and t^2)	5.85		
Quadratic and interaction terms			
$(OWN+EXT)^2$	93.14	65.04	59.73
$SPILL^2$	99.88	100.00	100.00
(OWN+EXT)×SPILL	79.47	100.00	100.00
(OWN+EXT)×CLIM	56.99*	0.00	0.00
SPILL×CLIM	45.27*	0.00	0.00

* Probabilities conditional on climate variable is added.

Table 2.4. Posterior inclusion probabilities of lag structures, BMA (%)

	Own Research			Own Extension			Spill-in research		
	All	Group A	Group B	All	Group A	Group B	All	Group A	Group B
By *b* parameters									
0.2-0.8	37.6	50.9	61.1	34.6	36.8	42.5	62.0		
1.0-1.4	25.4	23.1	20.3	26.1	21.4	24.9	26.4	26.3	
1.6-2.4	37.0	25.9	18.6	39.2	41.9	32.6	11.6	73.7	100.0
By peaks									
3	2.2	5.6	11.5						
4	12.4	67.8	88.2	0.1					
5	4.4	15.7	0.3	48.5	51.4	35.3			
6	3.5	2.9		32.2	23.5	26.3	0.1		
7	5.0	0.3		11.8	15.4	22.5			
8	6.4			7.4	9.7	15.9			
9-14	15.1						11.2		
15-18	17.1						86.1	100.0	100.0
19-20	13.4						2.0		
21-24	20.3	7.6					0.5		

Table 2.5. Summary statistics of lag structure parameters

	Own research		Own extension		Spill-in research	
	b	peak	*b*	peak	*b*	peak
Group A, BMA						
Mean	1.37	4.12	1.53	5.51	1.78	16.33
Median	1.40	4.00	1.60	5.00	1.60	16.00
Spatial median	1.35	4.00	1.57	5.00	1.74	16.00
SE	0.64	1.08	0.63	0.87	0.40	0.56
Group B, BMA						
Mean	1.33	3.92	1.38	5.70	2.25	16.75
Median	1.20	4.00	1.40	5.00	2.20	17.00
Spatial median	1.31	4.00	1.42	5.00	2.25	17.00
SE	0.70	0.28	0.65	1.00	0.17	0.61
BHM						
Mean	1.27	11.66	1.30	8.47	1.36	12.56
Median	1.27	6.30	1.30	6.29	1.37	10.33
Spatial median	1.28	7.63	1.30	6.92	1.37	11.31
SE	0.61	15.13	0.60	8.77	0.59	11.34

Table 2.6. Summary statistics of elasticities of MFP w.r.t. knowledge stocks

	Own research			Own extension			Spill-in research		
	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM
Mean	0.020	0.059	0.097	0.023	0.068	0.128	-0.050	-0.050	0.088
Median	0.022	0.058	0.106	0.025	0.066	0.134	-0.051	-0.050	0.095
Spatial median	0.022	0.058	0.105	0.025	0.066	0.136	-0.050	-0.050	0.100
SE	0.067	0.057	0.192	0.077	0.066	0.252	0.097	0.097	0.394
Empirical p-value	0.370	0.143	0.286	0.370	0.143	0.286	0.699	0.697	0.396
Quantiles									
0.005	-0.160	-0.088	-0.468	-0.186	-0.102	-0.622	-0.303	-0.299	-1.054
0.025	-0.114	-0.051	-0.303	-0.132	-0.058	-0.395	-0.241	-0.239	-0.721
0.050	-0.092	-0.032	-0.227	-0.106	-0.037	-0.292	-0.211	-0.208	-0.567
0.100	-0.066	-0.012	-0.146	-0.076	-0.014	-0.185	-0.175	-0.173	-0.402
0.250	-0.023	0.021	-0.021	-0.027	0.024	-0.026	-0.116	-0.115	-0.154
0.750	0.066	0.096	0.220	0.075	0.109	0.292	0.015	0.015	0.336
0.900	0.104	0.131	0.325	0.119	0.151	0.437	0.074	0.074	0.566
0.950	0.127	0.154	0.396	0.146	0.178	0.527	0.110	0.110	0.717
0.975	0.147	0.175	0.465	0.170	0.203	0.610	0.141	0.141	0.862
0.995	0.189	0.220	0.620	0.219	0.257	0.787	0.202	0.202	1.192

Table 2.7. Summary statistics of IRR (%)

	Own research			Own extension		
	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM
Mean	-16.1	20.9	46.1	-19.8	14.6	43.2
Median	12.9	38.0	19.8	11.7	31.2	43.1
Spatial median	0.3	35.8	26.4	-3.5	28.8	42.4
SE	67.3	54.4	202.4	63.7	50.7	182.9
Empirical p-value	0.434	0.194	0.304	0.434	0.194	0.304
Quantiles						
0.005	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.025	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.050	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.100	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.250	-100.0	12.4	-100.0	-100.0	10.8	-100.0
0.750	41.2	54.9	67.3	34.3	44.6	72.0
0.900	57.5	68.3	167.0	46.8	56.0	128.7
0.950	66.1	76.6	314.6	53.4	63.5	228.5
0.975	73.5	84.4	548.2	59.1	71.3	430.6
0.995	89.4	101.0	1284.1	71.7	91.1	1194.5

Table 2.8. Summary statistics of MIRR (%)

	Own research			Own extension		
	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM
Mean	-32.0	-6.5	-19.2	-25.8	3.8	-3.0
Median	5.2	9.3	11.9	8.9	21.2	31.1
Spatial median	-9.8	8.0	11.0	-2.3	20.5	32.0
SE	52.3	38.4	51.2	57.8	43.8	62.5
Empirical *p*-value	0.406	0.171	0.297	0.429	0.190	0.300
Quantiles						
0.005	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.025	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.050	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.100	-100.0	-100.0	-100.0	-100.0	-100.0	-100.0
0.250	-100.0	5.0	-100.0	-100.0	8.4	-100.0
0.750	9.8	11.5	15.2	23.0	28.2	42.9
0.900	11.9	12.9	17.0	29.4	32.9	49.5
0.950	12.8	13.6	18.0	32.4	35.4	52.7
0.975	13.4	14.2	18.8	34.6	37.4	55.2
0.995	14.6	15.3	20.2	38.5	41.1	59.7

Table 2.9. Summary statistics of growth accounting fractions (%)

	Own research			Own extension			Other		
	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM
Mean	4.0	15.3	8.9	1.9	7.1	4.4	94.2	77.6	86.7
Median	6.7	14.0	12.7	3.1	6.6	6.3	90.2	79.4	81.0
SE	774.1	167.8	69.3	365.0	80.1	35.2	1139.0	247.9	104.4
Quantiles									
0.005	-103.4	-32.4	-221.7	-48.1	-15.2	-114.6	16.6	2.7	-186.4
0.025	-40.5	-12.2	-142.1	-18.7	-5.7	-73.1	41.6	27.5	-106.4
0.050	-26.4	-6.1	-108.9	-12.2	-2.9	-55.7	51.8	37.9	-71.2
0.100	-15.4	-0.7	-75.6	-7.1	-0.3	-38.3	62.6	49.2	-34.0
0.250	-3.1	6.6	-29.5	-1.4	3.1	-14.7	77.1	66.0	22.1
0.750	15.6	23.2	51.8	7.3	10.8	25.9	104.6	90.4	144.2
0.900	25.5	34.7	89.0	11.9	16.1	45.0	122.5	101.0	213.9
0.950	32.8	42.4	113.9	15.3	19.7	57.6	138.6	109.0	264.5
0.975	39.8	49.5	137.8	18.6	23.0	69.3	159.2	117.9	315.0
0.995	56.9	66.4	192.1	26.5	30.9	95.7	251.6	147.6	435.5

Table 2.10. Summary statistics of simulations

	Own research			Own extension		
	Group A, BMA	Group B, BMA	BHM	Group A, BMA	Group B, BMA	BHM
Scenario 1: benefit/cost ratio for one million extra expenditure on own research/extension in 1992						
Mean	1.01	2.93	4.31	1.11	3.25	6.08
Median	1.09	2.90	4.46	1.21	3.17	6.30
Spatial median	1.09	2.89	4.47	1.20	3.17	6.07
SE	3.30	2.81	8.79	3.69	3.16	12.03
Empirical *p*-value	0.371	0.143	0.286	0.370	0.143	0.286
Scenario 2: benefit/cost ratio for one million extra expenditure on own research/extension in 1983						
Mean	1.30	3.74	5.41	1.16	3.35	6.38
Median	1.41	3.69	5.39	1.26	3.28	6.67
Spatial median	1.41	3.69	5.52	1.25	3.28	6.31
SE	4.21	3.57	11.31	3.80	3.24	12.57
Empirical *p*-value	0.369	0.143	0.286	0.369	0.143	0.286
Scenario 3: % change in 2016 MFP if the growth of own research/extension expenditure in 1992-2016 is same as in 1971-91						
Mean	0.83	2.38	3.01	1.34	3.92	7.02
Median	0.89	2.36	2.33	1.48	3.82	7.07
Spatial median	0.90	2.36	3.20	1.47	3.83	6.66
SE	2.68	2.26	6.75	4.50	3.82	13.84
Empirical *p*-value	0.370	0.144	0.286	0.370	0.144	0.286
Scenario 4: % change in 2016 MFP if the growth of own research/extension expenditure in 1992-2016 is same as in 1949-91						
Mean	2.31	6.62	8.21	3.50	10.23	18.66
Median	2.48	6.57	5.17	3.83	9.94	18.93
Spatial median	2.50	6.56	8.50	3.81	9.97	17.54
SE	7.43	6.29	19.10	11.70	9.99	37.15
Empirical *p*-value	0.370	0.143	0.286	0.370	0.143	0.286

Table 2.11. Summary statistics and test results of residuals of the first five years, BHM

	1949	1950	1951	1952	1953
Mean	-0.04725	-0.02557	0.02431	0.01184	-0.01813
SE	0.06190	0.05667	0.05200	0.04789	0.04437
Skewness	-0.072	-0.047	-0.027	-0.016	-0.010
Kurtosis	3.243	3.181	3.151	3.139	3.140
95% CI					
lower	-0.04727	-0.02559	0.02429	0.01182	-0.01814
upper	-0.04723	-0.02555	0.02432	0.01185	-0.01811
p-value of normality test	0.000	0.000	0.000	0.000	0.000
p-value of testing mean=0	0.000	0.000	0.000	0.000	0.000

Table 2.12. Tests and post-estimation results based on various central tendency measures of drawn parameters, BHM

	At the marginal posterior mean of parameters	At the marginal posterior median of parameters	At the spatial posterior median of parameters
p-values of tests			
Normality	0.676	0.262	0.205
Conditional mean	0.003	0.001	0.004
Conditional variance	0.001	0.001	0.180
Elasticities of MFP w.r.t. knowledge stocks			
Own research	0.179	0.115	0.103
Own extension	0.220	0.138	0.122
Spill-in research	-0.075	0.057	0.230
IRR (%)			
Own research	42.3	39.4	39.5
Own extension	55.5	45.0	37.5
MIRR (%)			
Own research	14.3	12.3	11.8
Own extension	38.3	31.4	29.4
Growth accounting fractions (%)			
Own research	16.1	15.3	8.5
Own extension	7.8	7.3	4.2
Other	76.1	77.4	87.3

Table 2.13. Percentage of models whose 95% CI of prediction contains the actual value, BMA

	With tests			Skip tests		
	All	Group A	Group B	All	Group A	Group B
2012	98.6	99.7	98.4	77.0	74.9	78.1
2013	99.9	99.9	99.2	81.6	79.9	83.3
2014	87.1	91.6	97.0	38.5	31.8	35.2
2015	80.9	86.7	98.9	40.8	33.7	37.1
2016	72.9	65.3	92.3	39.1	31.4	34.3
2012-16	72.6	64.9	90.1	33.6	27.1	29.8

Table 2.14. Relationship between rates of return and peaks of lag structures, BHM

Peak (year)	Percentage of draws (%)	IRR (%)		MIRR (%)	
		Own research	Own extension	Own research	Own extension
By peak of own research lag					
<=3	26.0	114.6	30.2	10.5	27.6
>3, <=5	15.8	60.3	46.2	12.0	32.3
>5, <=7	12.1	39.0	53.1	12.1	33.6
>7, <=14	22.0	14.2	32.2	7.3	25.3
>14, <=19	5.8	9.4	38.5	7.7	28.4
>19, <=24	4.3	13.1	51.4	11.6	37.1
>24	14.0	14.3	62.6	13.6	43.3
By peak of own extension lag					
<=4	26.0	26.7	86.6	10.0	30.4
>4, <=5	10.0	28.2	57.2	11.5	34.0
>5, <=6	10.8	27.9	50.5	11.9	34.8
>6, <=7	10.5	27.0	45.9	12.1	35.4
>7, <=8	8.2	29.1	38.9	11.8	33.9
>8	34.5	28.3	27.1	10.6	30.0
By peak of spill-in research lag					
<=9	44.2	28.2	56.3	13.6	39.4
>9, <=14	23.4	27.3	35.6	9.7	28.9
>14, <=20	20.5	-31.5	-34.4	-10.3	0.3
>20, <=24	4.0	20.6	26.8	7.7	23.2
>24	8.1	24.8	36.1	9.7	27.6

Figures

Figure 2.1. Overall workflow of BMA

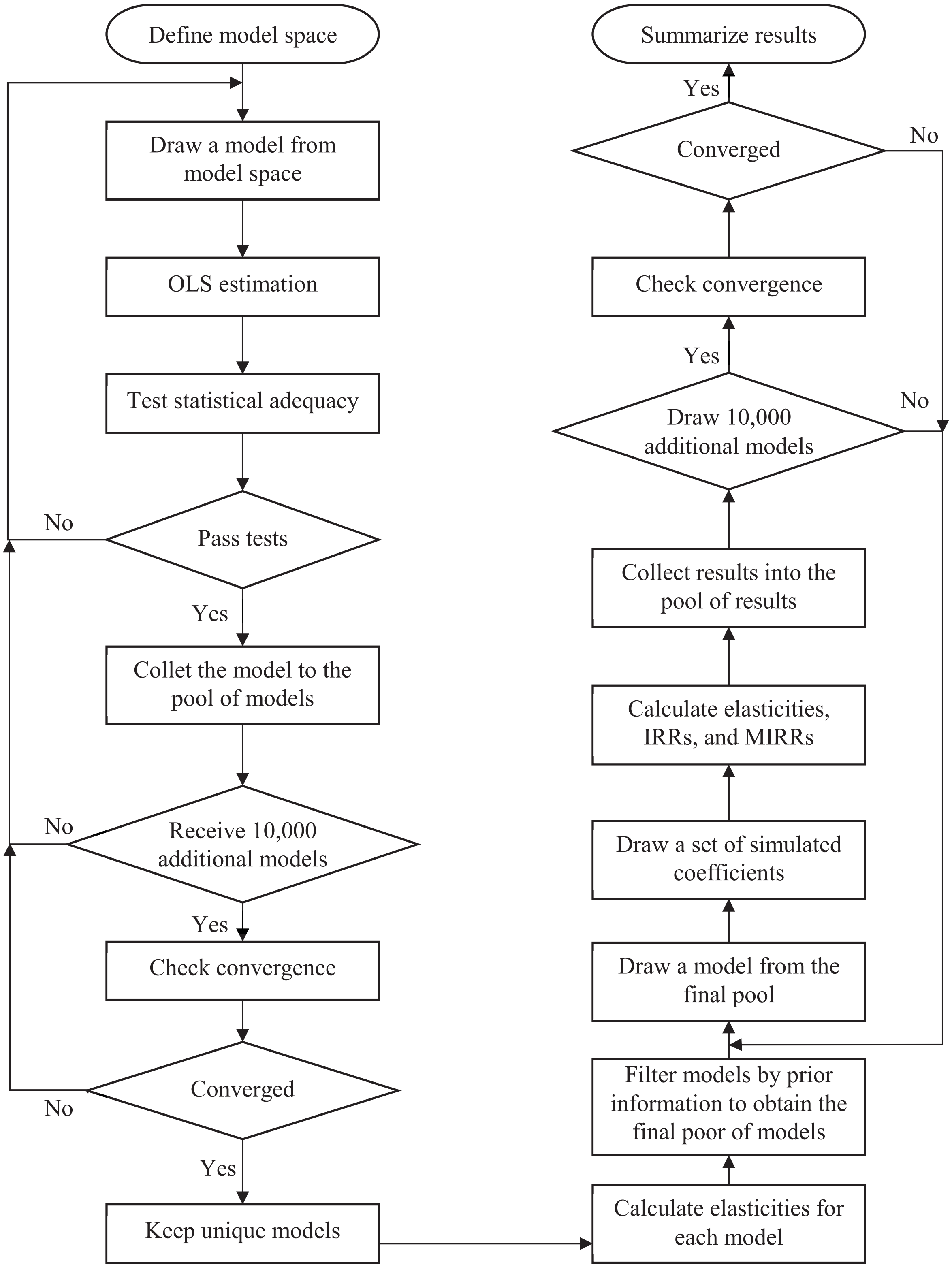


Figure 2.2. MFP of Virginia, neighbor states and US

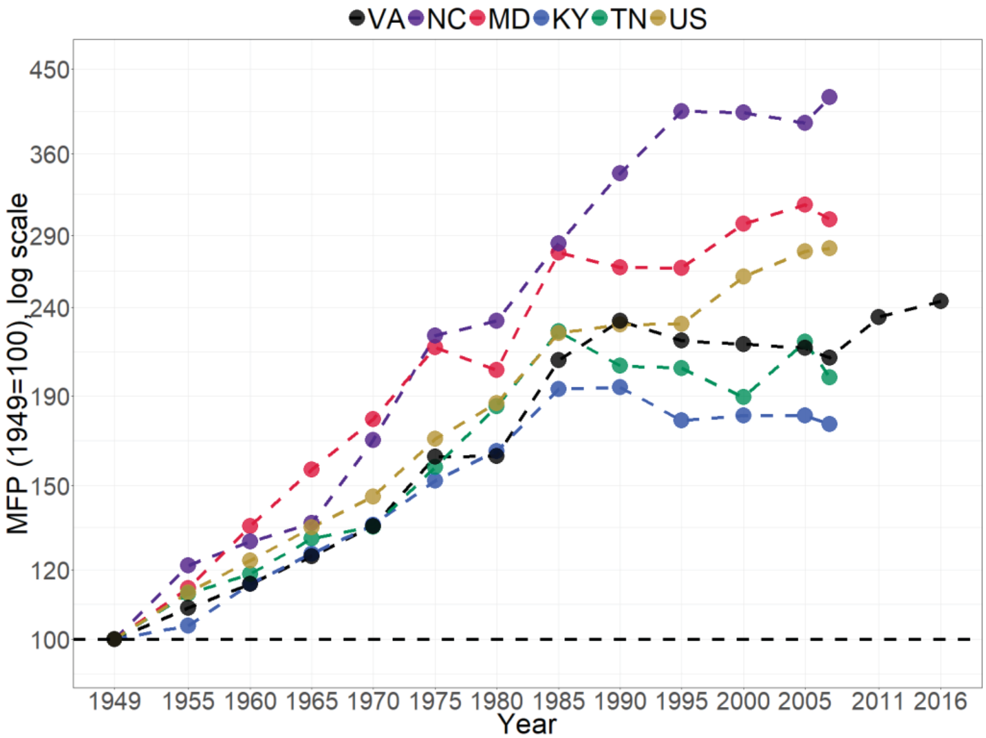


Figure 2.3. Components of Virginia MFP

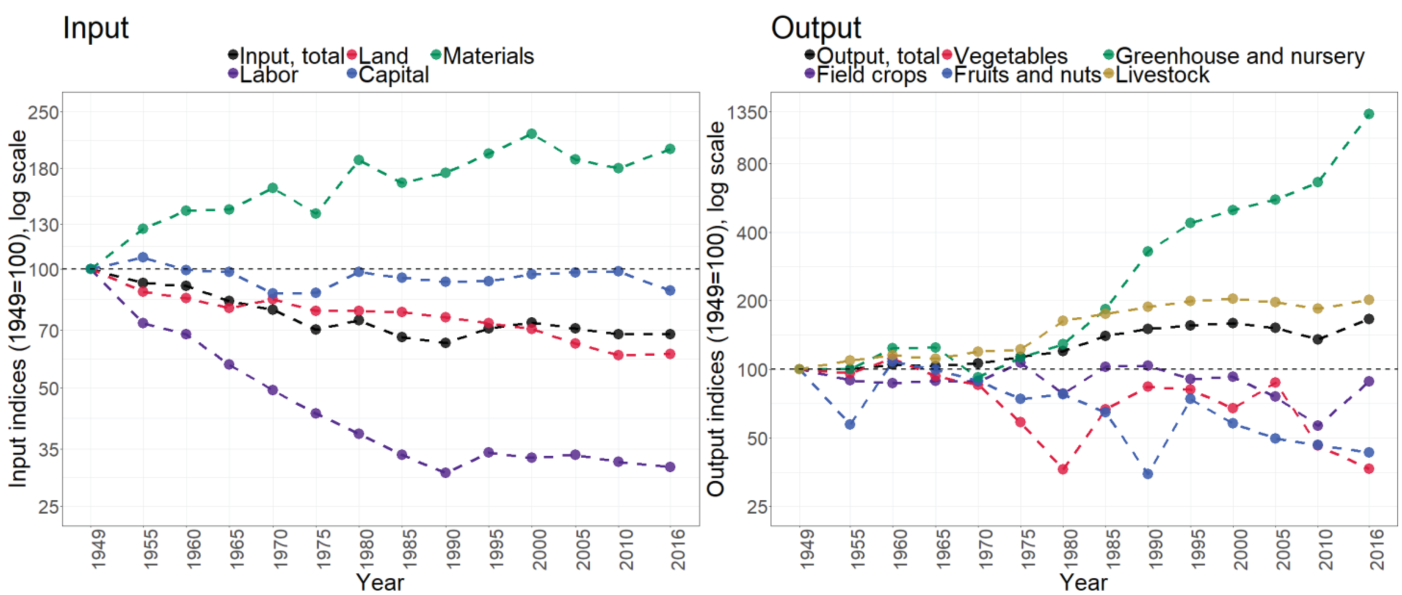


Figure 2.4. Agricultural R&E expenditures of Virginia and neighbor states

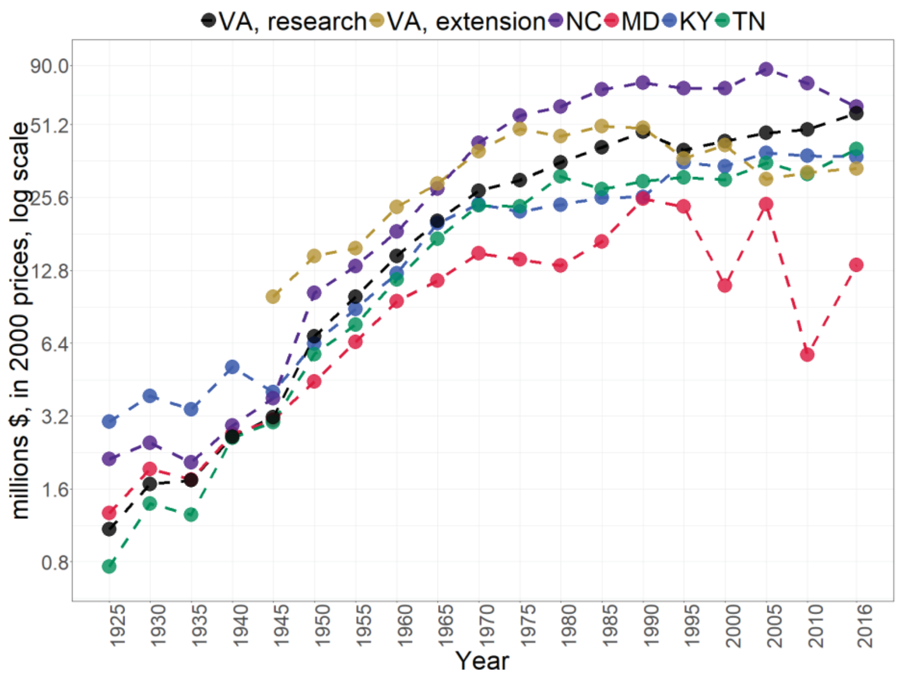
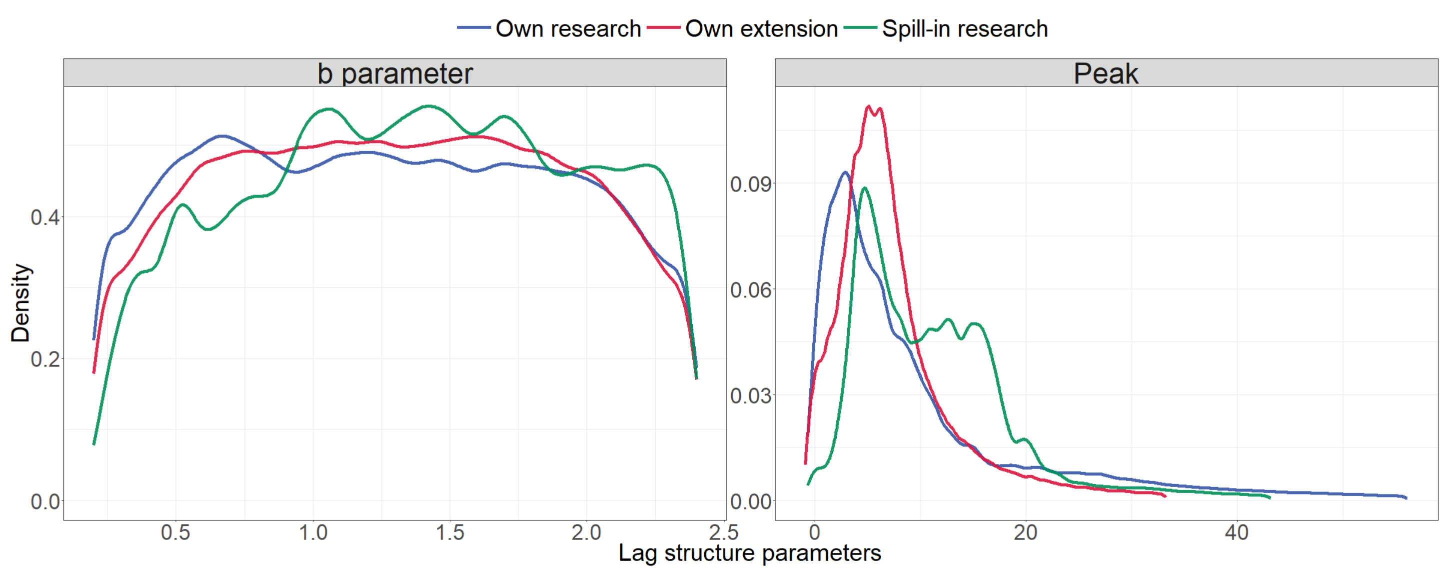


Figure 2.5. Distribution of lag structure parameters by BHM[44]



44 Since the distribution of the peaks has a long tail on the right tail, 2.5% of the draws are cut from the right tail when generating the density plot.

Figure 2.6. Pairwise scatter plots of lag structure parameters, BHM

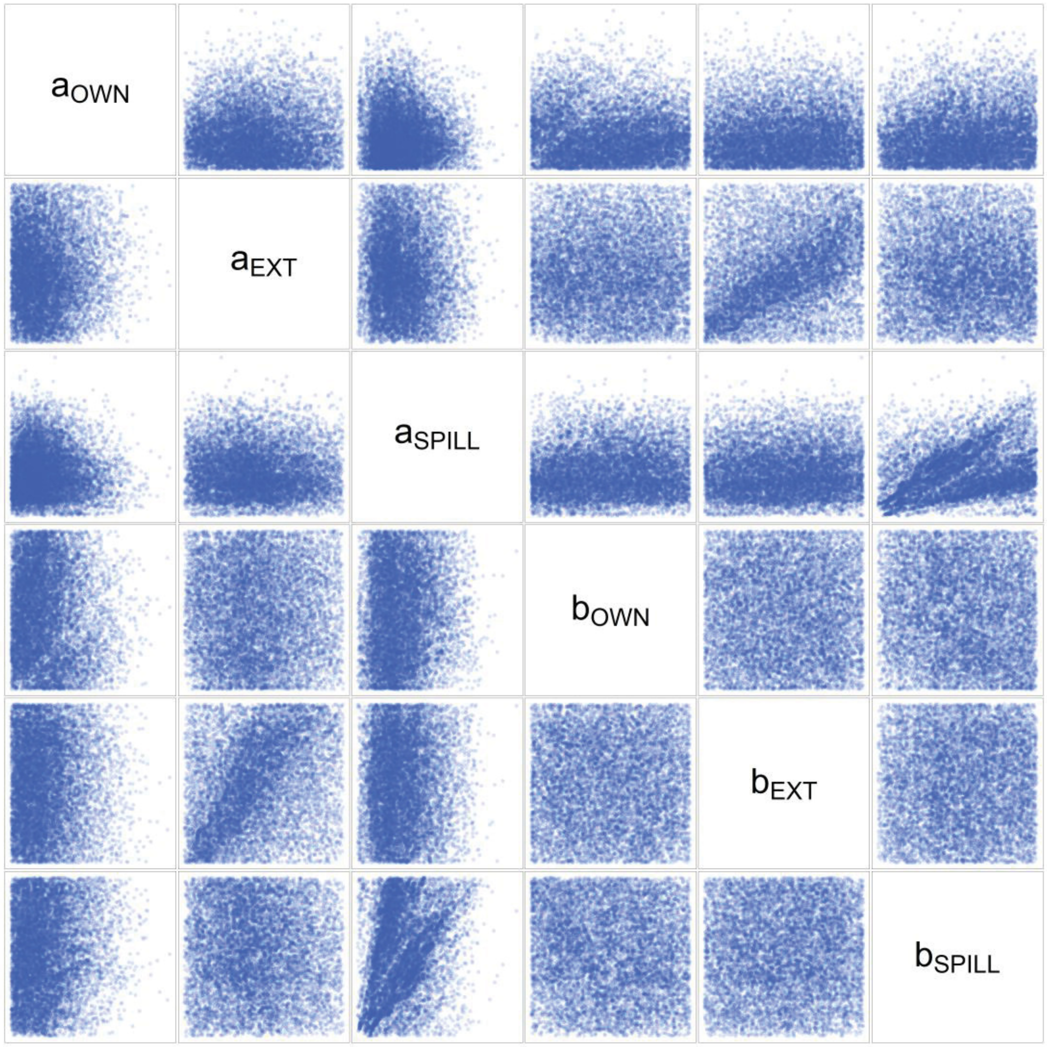
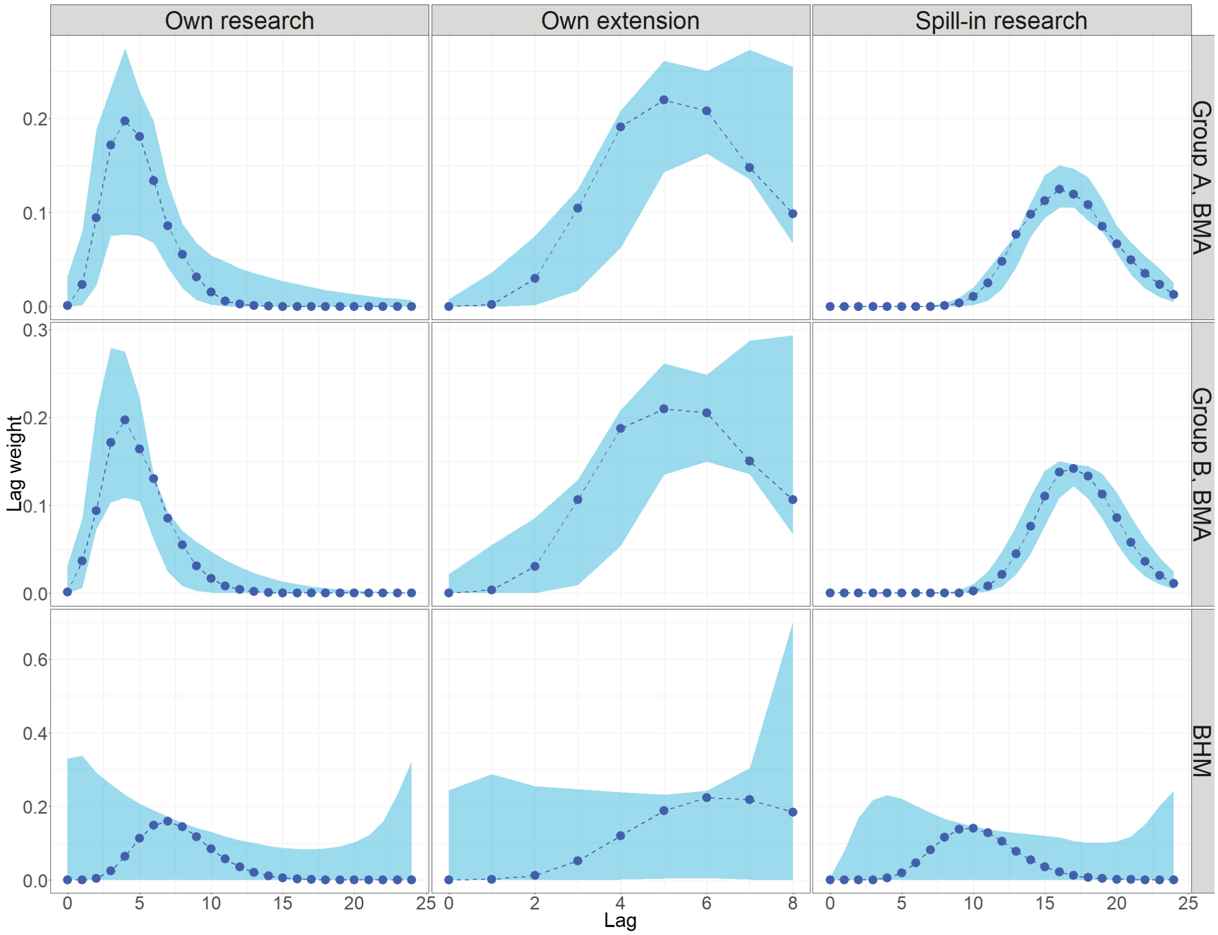


Figure 2.7. Typical lag structures and their uncertainty[45]



[45] The light blue area shows the 95% "confidence intervals" of lag weights.

Figure 2.8. Distributions of elasticities

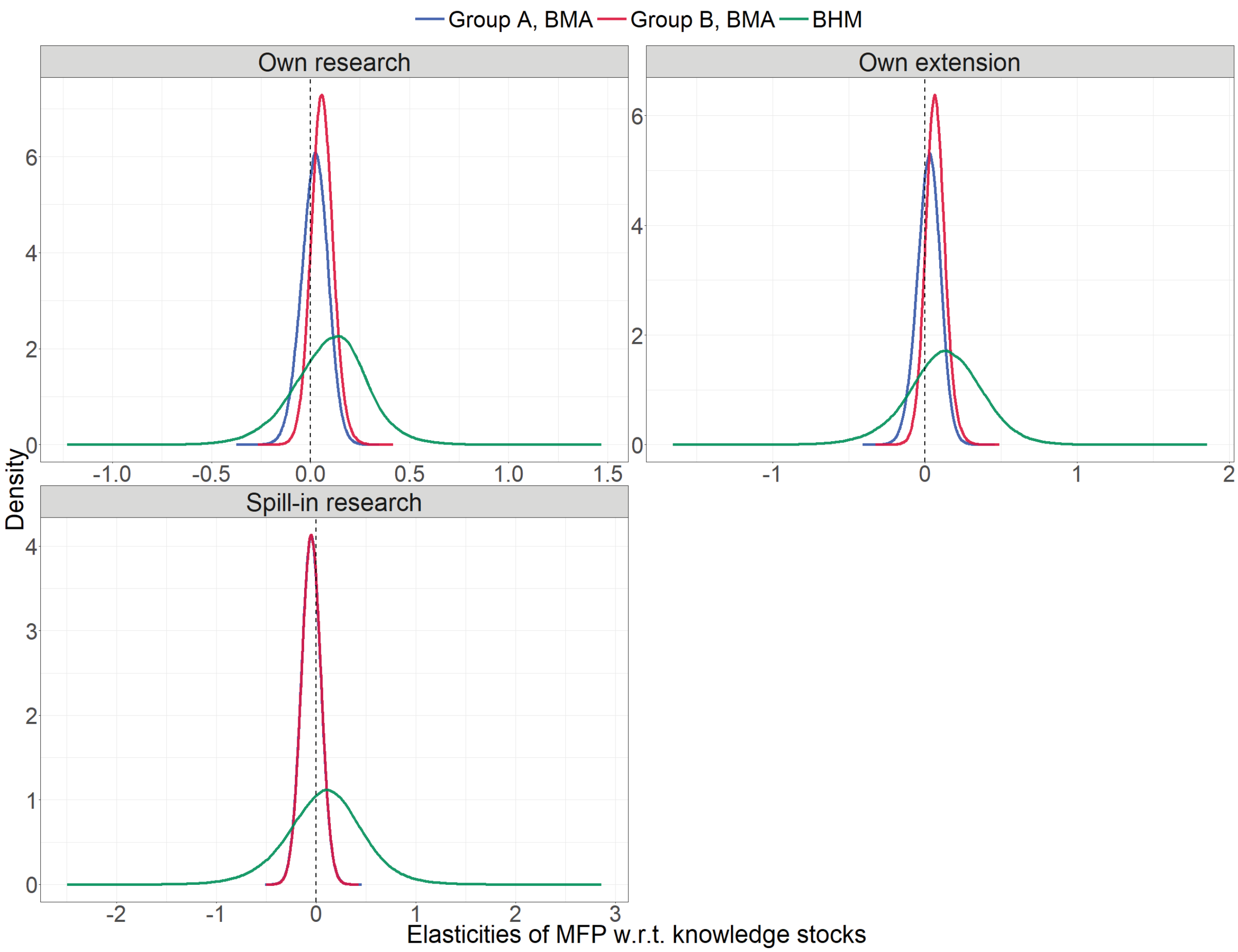
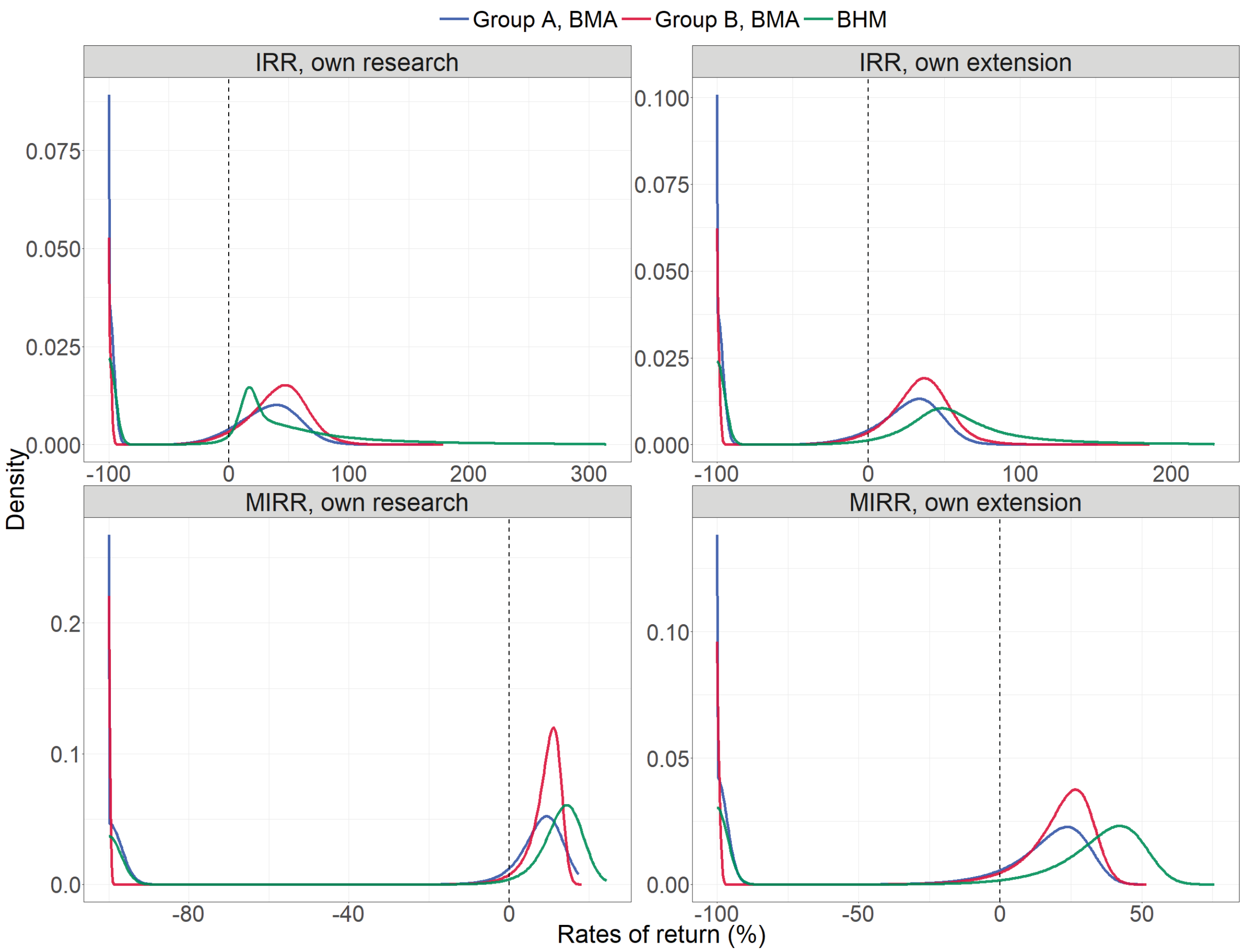
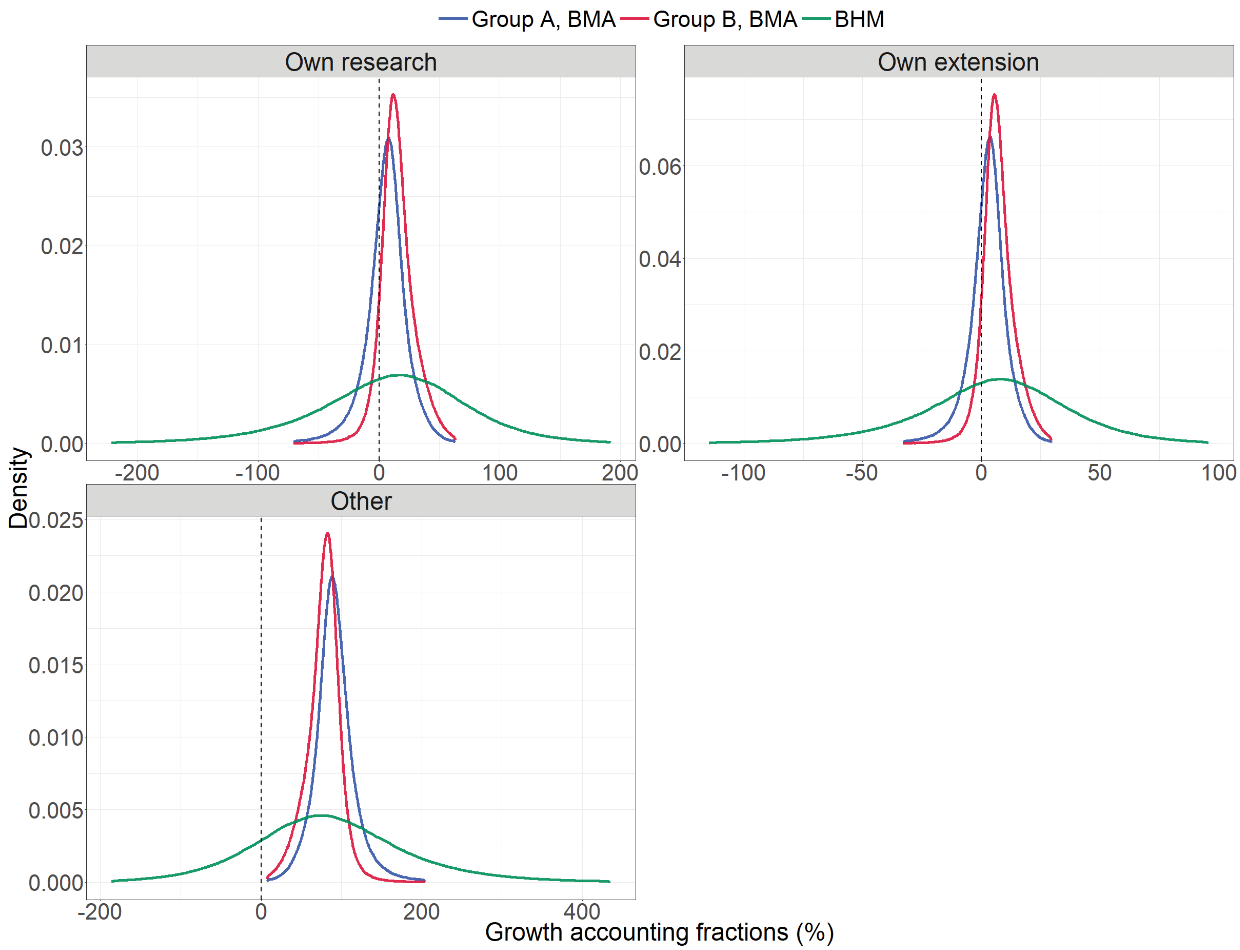


Figure 2.9. Distributions of rates of return[46]



46 Since the distributions of the IRRs from BHM have a very long and flat tail on the right end, 5% of the draws on the right tail are dropped.

Figure 2.10. Distributions of growth accounting fractions[47]



[47] Since the distributions have very long and flat tails at both tails, 0.5% of the draws are dropped from each tail.

Figure 2.11. Distributions of benefit/cost ratios and % change in 2016 MFP in simulated scenarios

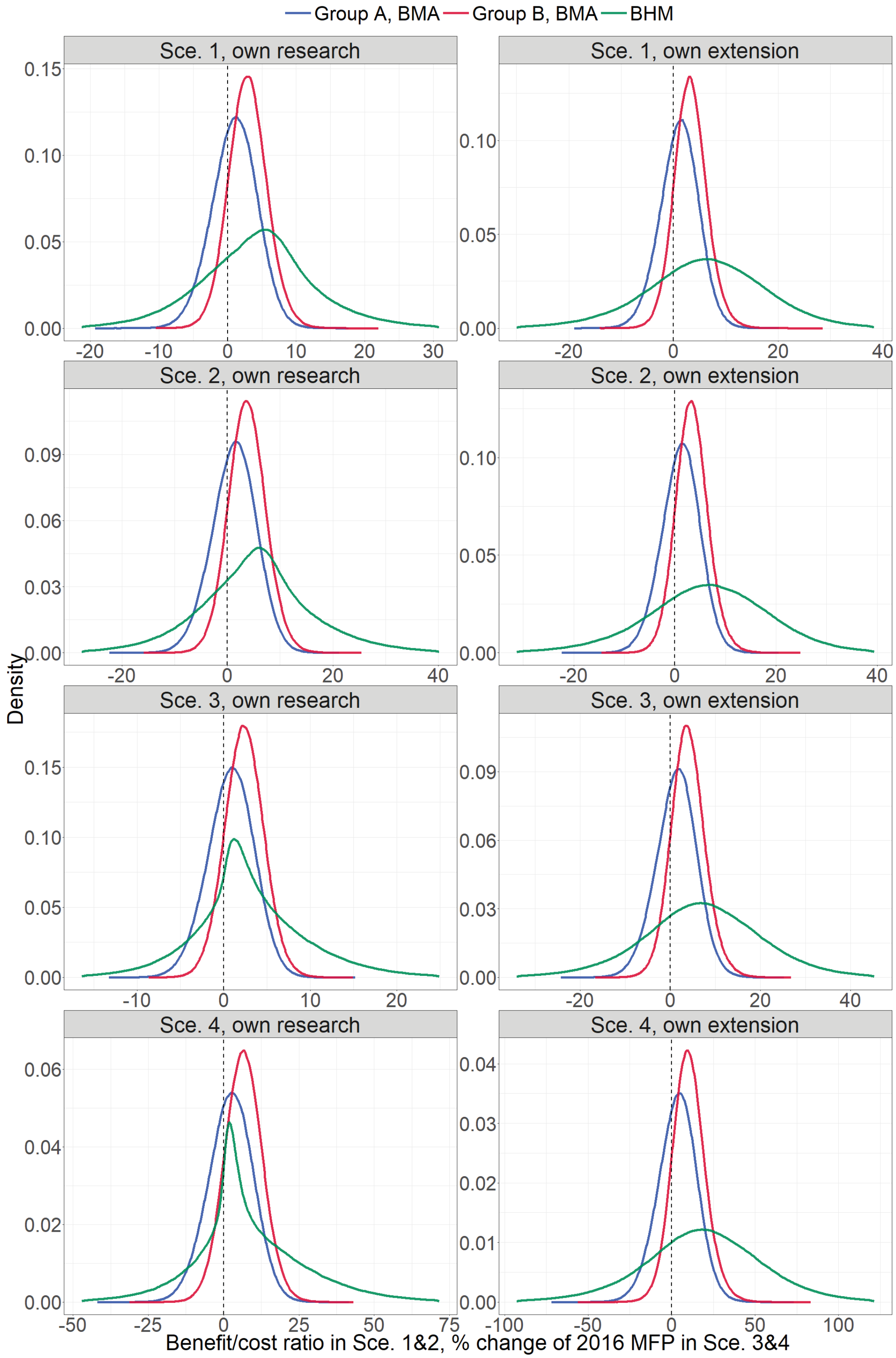
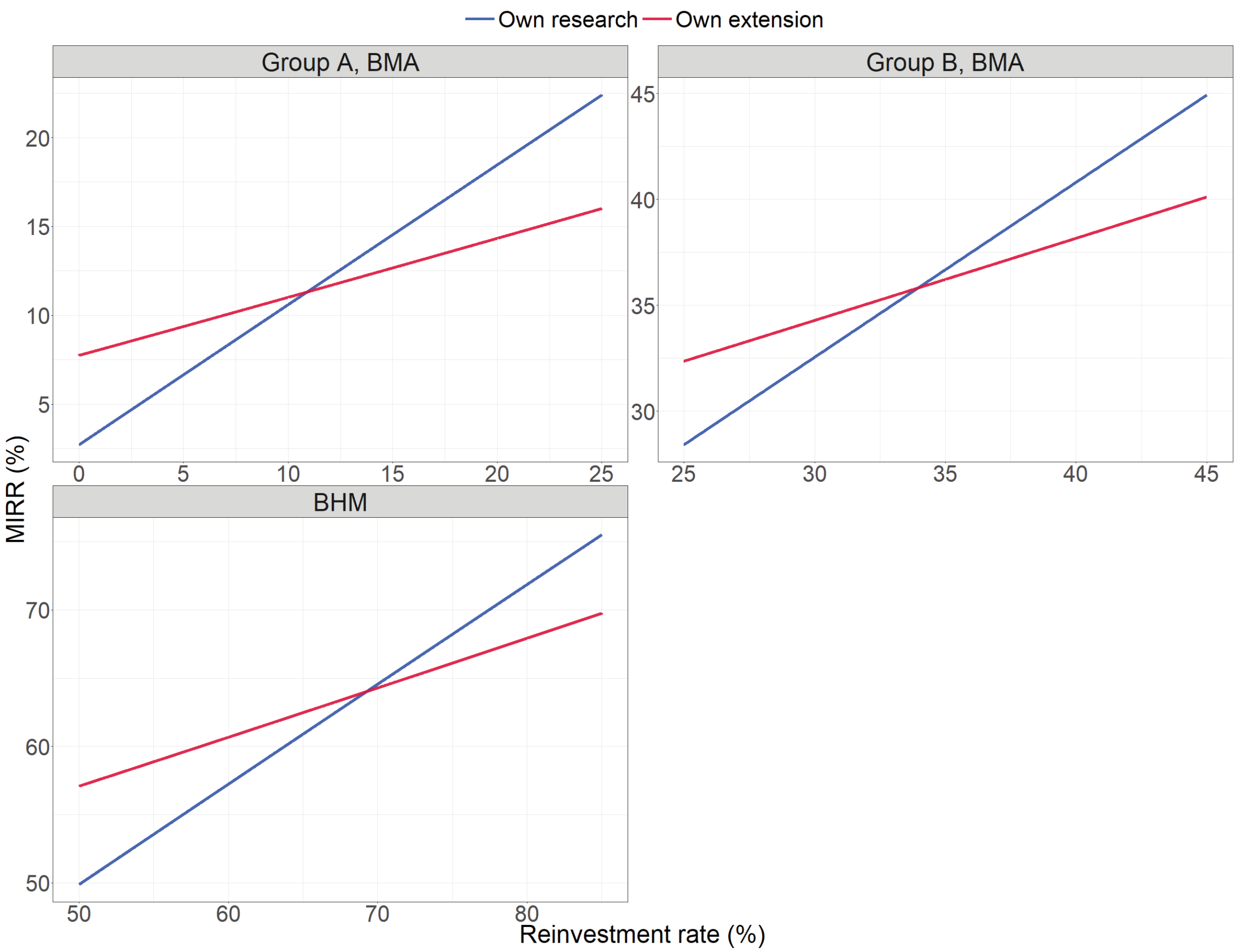


Figure 2.12. Relationship between MIRR and reinvestment rate



Appendix 2.1: Details about implementing Bayesian Averaging of Classical Estimates (BACE)

By Equation 2.2, we compare the posterior probabilities of two models M_0 and M_1,

$$\frac{P(M_1|D)}{P(M_0|D)} = \frac{P(D|M_1)}{P(D|M_0)} \times \frac{P(M_1)}{P(M_0)} \tag{2.A1}$$

Therefore, the ratio of posterior probabilities between two models (*posterior odds*) is the product of the ratio of model likelihoods (*Bayes factor*) and the ratio of model priors (*prior odds*) (Hoff, 2009). The difficulty of calculating posterior model probabilities (or posterior odds) is calculating the model likelihood (or Bayes factor) because it involves the integration over parameter space (Θ_i) given model:

$$P(D|M_i) = \int_{\Theta_i} P(D|\boldsymbol{\theta}_i, M_i) P(\boldsymbol{\theta}_i|M_i)\, \mathrm{d}\boldsymbol{\theta}_i \tag{2.A2}$$

Raftery (1995) derives an approximation of the model likelihood and relates the Bayes factor to the Bayesian Information Criterion (BIC):

$$2 \ln B_{10} \approx BIC_0 - BIC_1 \tag{2.A3}$$

B_{10} is the Bayes factor in Equation 2.A1. Therefore,

$$P(M_i|D) \propto \frac{P(M_i)}{\exp\left(\frac{1}{2} BIC_i\right)} \tag{2.A4}$$

For linear models with normal independent identically-distributed (NIID) error terms, Sala-i-Martin, Doppelhofer and Miller (2004) derive the following posterior model probabilities:

$$P(M_i|D) = \frac{P(M_i) N^{-\frac{k_i}{2}} SSE_i^{-\frac{N}{2}}}{\sum_j P(M_j) N^{-\frac{k_j}{2}} SSE_j^{-\frac{N}{2}}} \tag{2.A5}$$

N, k_i, and SSE_i are the number of observations, the number of covariates, and the estimated residual sum of squares. Equation 2.A3 is only an approximation of the Bayes factor; the error declines with sample size but does not vanish even with an infinite number of observations. However, the error tends to be very small if we assume the coefficients prior

$P(\boldsymbol{\theta}_i|M_i)$ is trivial (Raftery, 1995). That means the data is very informative, and likelihood $P(D|\boldsymbol{\theta}_i, M_i)$ dominate $P(\boldsymbol{\theta}_i|M_i)$ in model likelihood. In this study, the covariates have explained the majority of the variations in the agricultural productivity. Therefore, it is reasonable to expect the error between Equation 2.A5 and the true model posteriors is limited.

The two assumptions of Chipman (1996) are employed to reduce model priors to prior inclusion probabilities of covariates. The *conditional independence* assumption assumes that adding terms of a given order to a model are independent, conditional on all terms of a lower order. The *inheritance* assumption assumes that the probability of adding a higher-order term depends only on the probabilities of lower-order terms used to form it. By the assumptions, own research (γ_1), own extension (γ_2), spill-in research (γ_3), climate variable (γ_4), and time trends (γ_{10}) enter a model independently. Given knowledge stocks and climate variable, the probabilities of adding quadratic terms (γ_5, γ_6) and interactions (γ_7, γ_8, γ_9) depend only on the first-order terms used to form them. The prior probability of model $\boldsymbol{\gamma}$ is

$$\begin{aligned} P(\boldsymbol{\gamma}) = {} & P(\gamma_1)P(\gamma_2)P(\gamma_3)P(\gamma_4)P(\gamma_5|\gamma_1,\gamma_2)P(\gamma_6|\gamma_3)P(\gamma_7|\gamma_1,\gamma_2,\gamma_3) \\ & \times P(\gamma_8|\gamma_1,\gamma_2,\gamma_4)P(\gamma_9|\gamma_3,\gamma_4)P(\gamma_{10}) \end{aligned} \tag{2.A6}$$

For γ_i, $i \in \{4,5,6,7,10\}$, we assume the probabilities that they take each possible value are identical. For γ_i, $i \in \{8,9\}$, since we add the interactions only if climate variable is added ($\gamma_4 \neq 0$),

$$P(\gamma_i = 1) = \begin{cases} 0 & \text{if } \gamma_4 = 0 \\ 1/2 & \text{if } \gamma_4 \neq 0 \end{cases} \tag{2.A7}$$

As a result,

$$P(\boldsymbol{\gamma}) \propto P(\gamma_1)P(\gamma_2)P(\gamma_3)P(\gamma_8|\gamma_1,\gamma_2,\gamma_4)P(\gamma_9|\gamma_3,\gamma_4) \tag{2.A8}$$

As discussed in the paper, it is improper to use diffuse prior inclusion probabilities for γ_1, γ_2, and γ_3. The priors of knowledge stocks with higher similarities to other stocks require dilution. To understand the necessity, suppose we are dealing with a simple variable selection problem for a univariate CLRM $y = \alpha_0 + \alpha_1 x + u$. Initially, we choose between two alternative covariates $\{x_1, x_2\}$ and assign equal prior inclusion probability (1/2) to each of them. Then we add another variable x_3 that is exactly the same as x_2 to the set of alternative covariates. Since x_3 and x_2 are identical, this action should *dilute* the inclusion probability of x_2 to 1/4 but not influence that of x_1. This is similar to the car/red bus/blue bus argument usually used to criticize the Independence of Irrelevant Alternatives (IIA) assumption. Someone chooses between a red bus and a car, and a third choice of blue bus is offered to him/her; the extra choice should *dilute* the probability of choosing a red bus but has trivial effect on the probability of choosing a car.

George (2010) discusses several methods for diluting model priors, and the one that is relevant for us is *dilution by distances*: prior inclusion probability of a covariate is proportional to its distance to other covariates. In the context of OLS, a straightforward measure of similarity between two variables is the Pearson correlation coefficient: two variables with a correlation coefficient of one function identically in an OLS regression. Since we need a measure of distance and the correlation between knowledge stocks are always positive, a reasonable choice is 1-correlation coefficient. For each knowledge stock, let the prior inclusion probability of being:

$$P(\gamma_i = j) \propto \frac{\sum_{m \neq j}\left[1 - cor\left(S_{ij}, S_{im}\right)\right]}{n_i - 1}, i = 1,2,3 \qquad (2.A9)$$

$cor\left(S_{ij}, S_{im}\right)$ is the correlation coefficient between knowledge stocks S_{ij} and S_{im}; n_i is the number of different stocks in category i (1 for own research, 2 for own extension, and 3 for spill-in research). To summarize, to measure the distance between S_{ij} and other stocks in the same category, we calculate 1-correlation coefficient between S_{ij} and other stocks and average them. Since model priors is reduced to the product of prior inclusion probabilities of covariates, we draw covariates when we draw models the model space. To draw a model, we first independently draw γ_1, γ_2, γ_3, γ_4, and γ_{10}, and then determine whether each quadratic and interaction term is added to the model or not. As described in the paper, for drawing the covariates, diffuse sampling weights are used for initial draws, while posterior inclusion probabilities are used afterward.

Appendix 2.2: Details about implementing Bayesian Hierarchical Model (BHM)

This appendix presents the mathematical details of the Bayesian hierarchical model (BHM): the prior and the conditional posterior probabilities of parameters and the implementation of the Gibbs sampling (GS) and Metropolis-Hastings (M-H) algorithm. For briefness, the notations in this appendix are slightly different from the main body.

Notations

The top model of BHM links agricultural productivity with the knowledge stocks and control variables:

$$\begin{aligned} Y_t = \beta_0 + \beta_1(S_{1t} + S_{2t}) + \beta_2 S_{3t} + \beta_3(S_{1t} + S_{2t})^2 + \beta_4 S_{3t}^2 + \beta_5(S_{1t} + S_{2t}) \\ \times S_{3t} + \beta_6 t \ + \sigma_Y \varepsilon_{Yt}, \boldsymbol{\varepsilon}_Y \sim \mathrm{N}(\mathbf{0}, \mathbf{I}) \end{aligned} \tag{2.A10}$$

Y_t is the logarithm of MFP in year t. S_{1t}, S_{2t} and S_{3t} are own research, own extension and spill-in research knowledge stocks in year t respectively. $\mathbf{I}$ is a $n \times n$ identity matrix (n is sample size).

The bottom models link knowledge stocks with agricultural R&E expenditures:

$$S_{kt} = \sum_{j=0}^{T_k - 1} w_{kj} R_{k,t-j} + \sigma_{S_k} \varepsilon_{S_k,t}, \boldsymbol{\varepsilon}_{S_k} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}), k = 1,2,3 \tag{2.A11}$$

T_k is the lag length. w_{kj} is the lag weight of lag j ($j = 0,1,\cdots,T_k - 1$). $R_{k,t-j}$ is the research or extension expenditure in year $t - j$. The lag weights are assumed to follow the shape of Gamma distribution:

$$w_{kj} = \frac{(j+1)^{a_k - 1} e^{-b_k (j+1)}}{\sum_{i=0}^{T_k - 1} (i+1)^{a_k - 1} e^{-b_k (i+1)}} \tag{2.A12}$$

$a_k > 0$ and $b_k > 0$ are the shape and rate parameters of the Gamma distribution. For convenience, we define the following notations:

- $\mathbf{Y}$ ($n \times 1$): the vector of dependent variables in the top model
- $\mathbf{S}_k$ ($n \times 1$, $k = 1,2,3$): the vector of knowledge stocks ($k = 1$ for own research, $k = 2$ for own extension, $k = 3$ for spill-in research)
- $\mathbf{S} = \{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3\}$: the set of three knowledge stocks

- $\mathbf{S}_{-k}$: $\mathbf{S}$ excludes $\mathbf{S}_k$
- $\mathbf{w}_k$ ($T_k \times 1$, $k = 1,2,3$): the vector of lag weights
- $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \sigma_Y^2, \sigma_{S_1}^2, \sigma_{S_2}^2, \sigma_{S_3}^2, a_1, a_2, a_3, b_1, b_2, b_3\}$: the set of parameters to be estimated; $\boldsymbol{\beta} = \{\beta_0, \beta_1, \cdots, \beta_6\}$ is the vector of coefficients in the top model
- $\boldsymbol{\theta}_{-\theta_i} = \boldsymbol{\theta} \backslash \{\theta_i\}$: $\boldsymbol{\theta}$ excludes parameter θ_i
- $\mathbf{X} = [\mathbf{1} \quad \mathbf{S}_1 + \mathbf{S}_2 \quad \mathbf{S}_3 \quad (\mathbf{S}_1 + \mathbf{S}_2)^2 \quad \mathbf{S}_3^2 \quad (\mathbf{S}_1 + \mathbf{S}_2)\mathbf{S}_3 \quad \mathbf{t}]$ $(n \times 7)$: the matrix of covariates in the top model
- matrix $\mathbf{R}_k$ ($n \times T_k$, $k = 1,2,3$):

$$\mathbf{R}_k = \begin{bmatrix} R_{k,1} & R_{k,0} & \cdots & R_{k,2-T_k} \\ R_{k,2} & R_{k,1} & \cdots & R_{k,3-T_k} \\ \vdots & \vdots & \ddots & \vdots \\ R_{k,n-1} & R_{k,n-2} & \cdots & R_{k,n-T_k} \\ R_{k,n} & R_{k,n-1} & \cdots & R_{k,n-T_k+1} \end{bmatrix} \tag{2.A13}$$

  Therefore, $\mathbf{S}_k - \mathbf{R}_k \mathbf{w}_k = \sigma_{S_k} \boldsymbol{\varepsilon}_{S_k}$ is the vector of residuals in the bottom models.
- $\mathbf{D} = \{\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \mathbf{t}\}$: the set of "original" explanatory data

Priors of parameters

Following the conventions in the Bayesian modeling of classical linear regression models, we assume the normal prior for $\boldsymbol{\beta}$ and the inverse Gamma prior for the variances of error terms regardless of top or bottom models:

$$\boldsymbol{\beta} \sim \mathrm{N}(\boldsymbol{\mu}_0, \mathbf{V}_0) \tag{2.A14}$$

$$\sigma^2 \sim \mathrm{IG}(v_0, \tau_0) \tag{2.A15}$$

$\boldsymbol{\mu}_0$ and $\mathbf{V}_0$ are the mean vector and variance-covariance matrix of the normal distribution. v_0 and τ_0 are the shape and scale parameters of the inverse Gamma distribution. These priors are preferred because they are *conjugate* priors, meaning that the posterior probabilities also follow the same types of distribution.

Since the parameters of the Gamma distributed lags have to be positive, we assume their priors to be truncated normal with the lower bound equals zero:

$$a_k \sim \mathrm{TN}\left(\mu_{a_k}, v_{a_k}^2, 0\right), k = 1,2,3 \tag{2.A16}$$

$$b_k \sim \mathrm{TN}\left(\mu_{b_k}, v_{b_k}^2, 0\right), k = 1,2,3 \tag{2.A17}$$

Conditional posterior probabilities of parameters

In BHM, the knowledge stocks are taken as parameters that depend on the unknown parameters $\boldsymbol{\theta}$ and data. The joint posterior probability of $\boldsymbol{\theta}$ and knowledge stocks can be decomposed as:

$$\begin{aligned} \mathrm{P}(\boldsymbol{\theta}, \boldsymbol{S}|\mathbf{Y}, \mathbf{D}) &= \frac{\mathrm{P}(\boldsymbol{\theta}, \boldsymbol{S})\mathrm{P}(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{S}, \mathbf{D})}{\mathrm{P}(\mathbf{Y}|\mathbf{D})} \\ &\propto \mathrm{P}(\boldsymbol{\theta}, \boldsymbol{S})\mathrm{P}(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{S}, \mathbf{D}) \\ &= \mathrm{P}(\boldsymbol{\theta})\ \ \mathrm{P}(\boldsymbol{S}|\boldsymbol{\theta})\mathrm{P}(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{S}, \mathbf{D}) \end{aligned} \tag{2.A18}$$

The prior probability of each element in $\boldsymbol{\theta}$ is independent of each other:

$$\mathrm{P}(\boldsymbol{\theta}) = \mathrm{P}(\boldsymbol{\beta})\mathrm{P}(\sigma_Y^2)\prod_{k=1}^{3}\left[\mathrm{P}\left(\sigma_{S_k}^2\right)\mathrm{P}(a_k)\mathrm{P}(b_k)\right] \tag{2.A19}$$

The probability of $\mathbf{S}_k$ conditional on $\boldsymbol{\theta}$ is also independent of other knowledge stocks:

$$\mathrm{P}(\boldsymbol{S}|\boldsymbol{\theta}) = \prod_{k=1}^{3}\mathrm{P}(\mathbf{S}_k|\boldsymbol{\theta}) = \prod_{k=1}^{3}\mathrm{P}\left(\mathbf{S}_k|\sigma_{S_k}^2, a_k, b_k\right) \tag{2.A20}$$

Therefore,

$$\mathrm{P}(\boldsymbol{\theta})\ \ \mathrm{P}(\boldsymbol{S}|\boldsymbol{\theta}) = \mathrm{P}(\boldsymbol{\beta})\mathrm{P}(\sigma_Y^2)\prod_{k=1}^{3}\left[\mathrm{P}\left(\sigma_{S_k}^2\right)\mathrm{P}(a_k)\mathrm{P}(b_k)\mathrm{P}\left(\mathbf{S}_k|\sigma_{S_k}^2, a_k, b_k\right)\right] \tag{2.A21}$$

Since the error terms are assumed to be NIID,

$$\begin{aligned} \mathrm{P}(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{S}, \mathbf{D}) &= \mathrm{P}(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \sigma_Y^2) \\ &= (2\pi)^{-\frac{n}{2}}(\sigma_Y^2)^{-\frac{n}{2}}\exp\left[-\frac{1}{2\sigma_Y^2}(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})\right] \\ &= (2\pi)^{-\frac{n}{2}}(\sigma_Y^2)^{-\frac{n}{2}}\exp\left(-\frac{RSS_Y}{2\sigma_Y^2}\right) \end{aligned} \tag{2.A22}$$

$$\mathrm{P}\left(\mathbf{S}_k|\sigma_{S_k}^2, a_k, b_k\right) = (2\pi)^{-\frac{n}{2}}\left(\sigma_{S_k}^2\right)^{-\frac{n}{2}}\exp\left[-\frac{1}{2\sigma_{S_k}^2}(\mathbf{S}_k - \mathbf{R}_k\mathbf{w}_k)'(\mathbf{S}_k - \mathbf{R}_k\mathbf{w}_k)\right]$$
$$= (2\pi)^{-\frac{n}{2}}\left(\sigma_{S_k}^2\right)^{-\frac{n}{2}}\exp\left(-\frac{RSS_{S_k}}{2\sigma_{S_k}^2}\right) \tag{2.A23}$$

RSS_Y and RSS_{S_k} are the residual sum of squares in the top and bottom models. Each component in Equation 2.A9 is now known. To derive the conditional posterior probabilities of an element of $\boldsymbol{\theta}$ or knowledge stock, we just need to extract the components that include the parameter or knowledge stock from Equation 2.A9.

Conditional posterior of knowledge stocks

The conditional posterior kernel of knowledge stocks does not follow any known distribution (c is a constant):

$$\mathrm{P}(\mathbf{S}_k|\mathbf{Y}, \mathbf{D}, \mathbf{S}_{-k}, \boldsymbol{\theta}) \propto \mathrm{P}(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{S}, \mathbf{D})\mathrm{P}\left(\mathbf{S}_k|\sigma_{S_k}^2, a_k, b_k\right)$$
$$\propto \exp\left(-\frac{RSS_Y}{2\sigma_Y^2}\right)\exp\left(-\frac{RSS_{S_k}}{2\sigma_{S_k}^2}\right) \tag{2.A24}$$
$$= \exp\left[-\frac{1}{2}\left(\frac{RSS_Y}{\sigma_Y^2} + \frac{RSS_{S_k}}{\sigma_{S_k}^2}\right)\right]$$

$$\ln \mathrm{P}(\mathbf{S}_k|\mathbf{Y}, \mathbf{D}, \mathbf{S}_{-k}, \boldsymbol{\theta}) = c - \frac{1}{2}\left(\frac{RSS_Y}{\sigma_Y^2} + \frac{RSS_{S_k}}{\sigma_{S_k}^2}\right) \tag{2.A25}$$

Conditional posterior of $\boldsymbol{\beta}$

The conjugate normal prior makes the conditional posterior of $\boldsymbol{\beta}$ also normal:

$$\mathrm{P}\left(\boldsymbol{\beta}|\mathbf{Y}, \mathbf{D}, \boldsymbol{S}, \boldsymbol{\theta}_{-\boldsymbol{\beta}}\right) \propto \mathrm{P}(\mathbf{Y}|\boldsymbol{\theta}, \boldsymbol{S}, \mathbf{D})\mathrm{P}(\boldsymbol{\beta})$$
$$\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}'\left(\boldsymbol{V}_0^{-1} + \frac{\mathbf{X}'\mathbf{X}}{\sigma_Y^2}\right)\boldsymbol{\beta} - 2\boldsymbol{\beta}'\left(\boldsymbol{V}_0^{-1}\boldsymbol{\mu}_0 + \frac{\mathbf{X}'\mathbf{Y}}{\sigma_Y^2}\right)\right]\right\} \tag{2.A26}$$

$$\boldsymbol{\beta}|\mathbf{Y}, \mathbf{D}, \boldsymbol{S}, \boldsymbol{\theta}_{-\boldsymbol{\beta}} \sim \mathrm{N}(\boldsymbol{\mu}_1, \mathbf{V}_1) \tag{2.A27}$$

$$\boldsymbol{\mu}_1 = \mathbf{V}_1\left(\mathbf{V}_0^{-1}\boldsymbol{\mu}_0 + \frac{\mathbf{X}'\mathbf{Y}}{\sigma_Y^2}\right) \tag{2.A28}$$

$$\mathbf{V}_1 = \left(\mathbf{V}_0^{-1} + \frac{\mathbf{X}'\mathbf{X}}{\sigma_Y^2}\right)^{-1} \tag{2.A29}$$

If we assume diffuse priors $\boldsymbol{\mu}_0 = \mathbf{0}$, $\mathbf{V}_0 = \sigma_{Y0}^2\mathbf{I}$,

$$\boldsymbol{\mu}_1 = \frac{\mathbf{V}_1\mathbf{X}'\mathbf{Y}}{\sigma_Y^2} \tag{2.A30}$$

$$\boldsymbol{V}_1 = \left(\frac{\mathbf{I}}{\sigma_{Y0}^2} + \frac{\mathbf{X}'\mathbf{X}}{\sigma_Y^2}\right)^{-1} \tag{2.A31}$$

Conditional posterior of σ_Y^2

The conjugate inverse Gamma prior makes the conditional posterior of σ_Y^2 also inverse Gamma:

$$\begin{aligned} \mathrm{P}\left(\sigma_Y^2|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-\sigma_Y^2}\right) &\propto \mathrm{P}(\mathbf{Y}|\boldsymbol{\theta},\boldsymbol{S},\mathbf{D})\mathrm{P}(\sigma_Y^2) \\ &\propto (\sigma_Y^2)^{-\frac{n}{2}-v_0-1}\exp\left[-\frac{1}{2\sigma_Y^2}(2\tau_0 + RSS_Y)\right] \end{aligned} \tag{2.A32}$$

$$\sigma_Y^2|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-\sigma_Y^2} \sim \mathrm{IG}(v_{1Y},\tau_{1Y}) \tag{2.A33}$$

$$v_{1Y} = v_0 + \frac{n}{2} \tag{2.A34}$$

$$\tau_{1Y} = \tau_0 + \frac{RSS_Y}{2} \tag{2.A35}$$

Conditional posterior of $\sigma_{S_k}^2$

Similar to σ_Y^2, the conditional posterior of $\sigma_{S_k}^2$ is also inverse Gamma:

$$\sigma_{S_k}^2|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-\sigma_{S_k}^2} \sim \mathrm{IG}\left(v_{1S_k},\tau_{1S_k}\right) \tag{2.A36}$$

$$v_{1S_k} = v_0 + \frac{n}{2} \tag{2.A37}$$

$$\tau_{1S_k} = \tau_0 + \frac{RSS_{S_k}}{2} \tag{2.A38}$$

Conditional posterior of a_k *and* b_k

The conditional posterior kernels of a_k and b_k do not follow any known distribution:

$$P(a_k|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-a_k}) \propto P(\mathbf{S}_k|\sigma_{S_k}^2, a_k, b_k)P(a_k)$$
$$\propto \exp\left\{-\frac{1}{2}\left[\frac{RSS_{S_k}}{\sigma_{S_k}^2} + \frac{(a_k-\mu_{a_k})^2}{v_{a_k}^2}\right]\right\} \tag{2.A39}$$
$$P(b_k|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-b_k}) \propto P(\mathbf{S}_k|\sigma_{S_k}^2, a_k, b_k)P(b_k)$$
$$\propto \exp\left\{-\frac{1}{2}\left[\frac{RSS_{S_k}}{\sigma_{S_k}^2} + \frac{(b_k-\mu_{b_k})^2}{v_{b_k}^2}\right]\right\} \tag{2.A40}$$

Therefore (c is a constant),

$$\ln P(a_k|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-a_k}) = c - \frac{1}{2}\left[\frac{RSS_{S_k}}{\sigma_{S_k}^2} + \frac{(a_k-\mu_{a_k})^2}{v_{a_k}^2}\right] \tag{2.A41}$$

$$\ln P(b_k|\mathbf{Y},\mathbf{D},\boldsymbol{S},\boldsymbol{\theta}_{-b_k}) = c - \frac{1}{2}\left[\frac{RSS_{S_k}}{\sigma_{S_k}^2} + \frac{(b_k-\mu_{b_k})^2}{v_{b_k}^2}\right] \tag{2.A42}$$

Dividing the parameter set for Gibbs Sampling

For using the Gibbs sampling method to obtain the posterior probabilities, the set of unknown parameters $\boldsymbol{\theta}$ and knowledge stocks $\boldsymbol{S}$ is partitioned into several subsets, and the following sequence is employed in drawing from the conditional posterior distributions: $\boldsymbol{\beta} \to \sigma_Y^2 \to \mathbf{S}_k \to a_k \to b_k \to \sigma_{S_k}^2$. $\mathbf{S}_k \to a_k \to b_k \to \sigma_{S_k}^2$ are three parallel paths. For $\boldsymbol{\beta}$, σ_Y^2 and $\sigma_{S_k}^2$, we just need to draw from the normal or inverse Gamma conditional posterior distributions. For $\mathbf{S}_k$, a_k and b_k, we have to resort to the Metropolis-Hastings (M-H) algorithm because their conditional posterior kernels does not follow any known probability distribution.

M-H of $\boldsymbol{S}_k$

We adopt the truncated normal distribution with lower bound of zero as the proposal distribution:

$$\mathbf{S}_k^c \sim TN\left(\mathbf{S}_k^o, \sigma_{S_{kq}}^2\mathbf{I}, \mathbf{0}\right) \tag{2.A43}$$

$\mathbf{S}_k^c$ and $\mathbf{S}_k^o$ are the new draw and the previous draw. Since the new draw depends on the previous draw, we adopt a random-walk chain to draw $\boldsymbol{S}_k$'s. We need to notice that the truncated normal distribution is not symmetric and thus cannot be neglected from the probability of acceptance:

$$\mathrm{P}(\mathbf{S}_k^c|\mathbf{S}_k^o) = \frac{(2\pi)^{-\frac{n}{2}}\left(\sigma_{S_{kq}}^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma_{S_{kq}}^2}(\mathbf{S}_k^c - \mathbf{S}_k^o)'(\mathbf{S}_k^c - \mathbf{S}_k^o)\right]}{1 - \Phi\left(-\frac{\mathbf{S}_k^o}{\sigma_{S_{kq}}}\right)} \tag{2.A44}$$

$$\mathrm{P}(\mathbf{S}_k^o|\mathbf{S}_k^c) = \frac{(2\pi)^{-\frac{n}{2}}\left(\sigma_{S_{kq}}^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma_{S_{kq}}^2}(\mathbf{S}_k^o - \mathbf{S}_k^c)'(\mathbf{S}_k^o - \mathbf{S}_k^c)\right]}{1 - \Phi\left(-\frac{\mathbf{S}_k^c}{\sigma_{S_{kq}}}\right)} \tag{2.A45}$$

The log-acceptance probability is:

$$\begin{aligned} &\ln \frac{\mathrm{P}(\mathbf{S}_k^c|\cdots)\mathrm{P}(\mathbf{S}_k^o|\mathbf{S}_k^c)}{\mathrm{P}(\mathbf{S}_k^o|\cdots)\mathrm{P}(\mathbf{S}_k^c|\mathbf{S}_k^o)} \\ &\qquad = [\ln \mathrm{P}(\mathbf{S}_k^c|\cdots) - \ln \mathrm{P}(\mathbf{S}_k^o|\cdots)] + [\ln \mathrm{P}(\mathbf{S}_k^o|\mathbf{S}_k^c) - \ln \mathrm{P}(\mathbf{S}_k^c|\mathbf{S}_k^o)] \end{aligned} \tag{2.A46}$$

$$\ln \mathrm{P}(\mathbf{S}_k^c|\cdots) - \ln \mathrm{P}(\mathbf{S}_k^o|\cdots) = -\frac{1}{2}\left(\frac{RSS_Y^c - RSS_Y^o}{\sigma_Y^2} + \frac{RSS_{Sk}^c - RSS_{Sk}^o}{\sigma_{S_k}^2}\right) \tag{2.A47}$$

$$\ln \mathrm{P}(\mathbf{S}_k^o|\mathbf{S}_k^c) - \ln \mathrm{P}(\mathbf{S}_k^c|\mathbf{S}_k^o) = -\ln\left[1 - \Phi\left(-\frac{\mathbf{S}_k^c}{\sigma_{S_{kq}}}\right)\right] + \ln\left[1 - \Phi\left(-\frac{\mathbf{S}_k^o}{\sigma_{S_{kq}}}\right)\right] \tag{2.A48}$$

For conciseness, the conditional posterior $\mathrm{P}(\boldsymbol{S}_k|\mathbf{Y}, \mathbf{D}, \mathbf{S}_{-k}, \boldsymbol{\theta})$ is abbreviated as $\mathrm{P}(\boldsymbol{S}_k|\cdots)$.

M-H of a_k *and* b_k

The implementation of M-H for a_k and for b_k are almost identical, so we only present that for a_k. We adopt the truncated normal distribution with lower bound $a_{k,lwr}$ and upper bound $a_{k,upr}$ as the proposal distribution:

$$a_k^c \sim \mathrm{TN}\left(a_k^o, \sigma_{a_{kq}}^2, a_{k,lwr}, a_{k,upr}\right) \tag{2.A49}$$

Similar to the knowledge stocks, we adopt a random-walk chain, and the truncated normal distribution is asymmetric:

$$P(a_k^c|a_k^o) = \frac{(2\pi)^{-\frac{n}{2}}\left(\sigma_{a_{kq}}^2\right)^{-\frac{n}{2}}\exp\left[-\frac{1}{2\sigma_{a_{kq}}^2}(a_k^c - a_k^o)^2\right]}{\Phi\left(\frac{a_{k,upr} - a_k^o}{\sigma_{a_{kq}}}\right) - \Phi\left(\frac{a_{k,lwr} - a_k^o}{\sigma_{a_{kq}}}\right)} \tag{2.A50}$$

$$P(a_k^o|a_k^c) = \frac{(2\pi)^{-\frac{n}{2}}\left(\sigma_{a_{kq}}^2\right)^{-\frac{n}{2}}\exp\left[-\frac{1}{2\sigma_{a_{kq}}^2}(a_k^o - a_k^c)^2\right]}{\Phi\left(\frac{a_{k,upr} - a_k^c}{\sigma_{a_{kq}}}\right) - \Phi\left(\frac{a_{k,lwr} - a_k^c}{\sigma_{a_{kq}}}\right)} \tag{2.A51}$$

The log-acceptance probability is:

$$\begin{aligned}\ln\frac{P(a_k^c|\cdots)P(a_k^o|a_k^c)}{P(a_k^o|\cdots)P(a_k^c|a_k^o)} \\ &= [\ln P(a_k^c|\cdots) - \ln P(a_k^o|\cdots)] + [\ln P(a_k^o|a_k^c) - \ln P(a_k^c|a_k^o)]\end{aligned} \tag{2.A52}$$

$$\begin{aligned}\ln P(a_k^c|\cdots) - \ln P(a_k^o|\cdots) \\ &= -\frac{1}{2}\left[\frac{RSS_{S_k}^c - RSS_{S_k}^o}{\sigma_{S_k}^2} + \frac{\left(a_k^c - \mu_{a_k}\right)^2 - \left(a_k^o - \mu_{a_k}\right)^2}{v_{a_k}^2}\right]\end{aligned} \tag{2.A53}$$

$$\begin{aligned}\ln P(a_k^o|a_k^c) - \ln P(a_k^c|a_k^o) \\ &= -\ln\left[\Phi\left(\frac{a_{k,upr} - a_k^c}{\sigma_{a_{kq}}}\right) - \Phi\left(\frac{a_{k,lwr} - a_k^c}{\sigma_{a_{kq}}}\right)\right] \\ &\quad + \ln\left[\Phi\left(\frac{a_{k,upr} - a_k^o}{\sigma_{a_{kq}}}\right) - \Phi\left(\frac{a_{k,lwr} - a_k^o}{\sigma_{a_{kq}}}\right)\right]\end{aligned} \tag{2.A54}$$

Similarly, the conditional posterior $P\left(a_k|\mathbf{Y}, \mathbf{D}, \boldsymbol{S}, \boldsymbol{\theta}_{-a_k}\right)$ is abbreviated as $P(a_k|\cdots)$.

Appendix Tables

App. Table 2.5. Posterior inclusion probabilities of lags if SA tests are skipped in BMA (%)

	Own research		Own extension		Spill-in research	
	Group A	Group B	Group A	Group B	Group A	Group B
By b parameters						
0.2	8.5	8.5	7.9	7.8	8.0	7.8
0.4	8.4	8.3	8.1	8.1	8.1	7.9
0.6	8.3	8.2	8.2	8.2	8.2	8.0
0.8	8.3	8.3	8.3	8.3	8.3	8.2
1.0	8.3	8.3	8.3	8.3	8.3	8.3
1.2	8.3	8.3	8.4	8.4	8.4	8.4
1.4	8.3	8.4	8.4	8.4	8.4	8.4
1.6	8.3	8.3	8.5	8.5	8.4	8.5
1.8	8.3	8.3	8.5	8.5	8.4	8.5
2.0	8.3	8.4	8.5	8.5	8.5	8.6
2.2	8.4	8.4	8.5	8.5	8.5	8.6
2.4	8.3	8.4	8.6	8.6	8.5	8.7
By peaks						
4	5.3	6.4	16.3	15.5	3.1	3.9
5	5.3	6.6	16.6	16.2	3.2	3.7
6	5.1	6.6	16.7	16.8	3.3	3.5
7	5.0	6.4	16.7	17.1	3.3	3.4
8	4.9	6.2	16.8	17.2	3.3	3.4
9	4.9	5.9	16.9	17.2	3.3	3.4
10	4.9	5.6			3.5	3.6
11	4.8	5.0			3.7	4.0
12	4.6	4.6			4.2	4.5
13	4.5	4.1			4.8	5.3
14	4.3	3.8			5.0	5.5
15	4.2	3.6			5.0	5.2
16	4.1	3.5			4.9	4.9
17	4.1	3.4			5.0	4.6
18	4.1	3.4			5.1	4.5
19	4.1	3.4			5.3	4.6
20	4.2	3.4			5.4	4.7
21	4.3	3.4			5.5	4.9
22	4.3	3.5			5.6	5.2
23	4.3	3.6			5.7	5.5
24	4.4	3.7			5.8	5.8
25	4.4	3.8			5.8	5.9

App. Table 2.6. Posterior inclusion probabilities of covariates if SA tests are skipped in BMA (%)

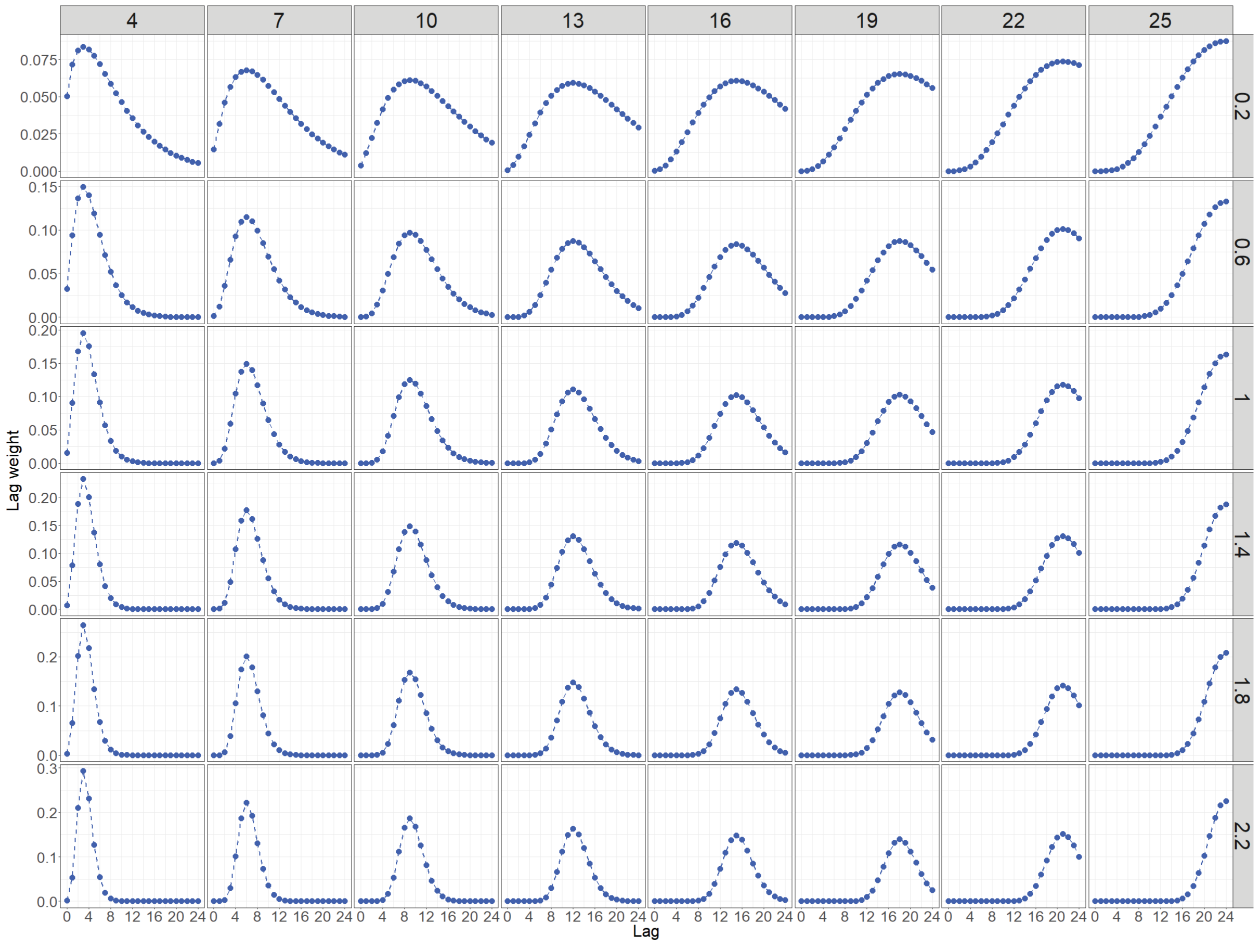
	All	Group A	Group B
States for calculating spill-in research			
Neighbor states	33.3	35.6	35.0
Similar by agro-ecological conditions	33.3	31.9	31.1
All states except VA	33.3	32.4	33.9
Climate variable			
Not included	33.3	33.3	32.8
Precipitation	33.4	33.4	34.3
PDSI	33.3	33.3	32.8
Time trend			
Not included	33.3	41.7	44.0
Linear (only t)	33.3	30.2	30.4
Quadratic (t, t^2)	33.4	28.0	25.6
Quadratic and interaction terms			
$(OWN+EXT)^2$	50.0	48.4	43.9
$SPILL^2$	50.0	48.4	42.4
(OWN+EXT)×SPILL	50.0	47.3	42.3
(OWN+EXT)×CLIM	50.0	50.0	50.1
SPILL×CLIM	50.0	50.0	50.0

App. Table 2.7. Summary statistics of elasticities and rates of return if SA tests are skipped in BMA

		Own research		Own extension		Spill-in research	
		Group A	Group B	Group A	Group B	Group A	Group B
Elasticity							
	Mean	0.205	0.283	0.242	0.336	0.608	0.475
	Median	0.187	0.215	0.217	0.249	0.542	0.446
	SE	0.209	0.191	0.254	0.232	0.475	0.335
	Empirical *p*-value	0.149	0.010	0.149	0.010	0.077	0.060
	Quantiles						
	0.025	-0.074	0.034	-0.119	0.043	-0.105	-0.094
	0.050	-0.050	0.070	-0.080	0.088	-0.047	-0.024
	0.950	0.659	0.698	0.793	0.840	1.446	1.057
	0.975	0.738	0.771	0.888	0.928	1.572	1.174
IRR (%)							
	Mean	40.9	73.5	96.4	146.7		
	Median	62.6	73.3	61.2	86.2		
	SE	67.1	32.4	156.6	151.1		
	Empirical *p*-value	0.169	0.014	0.169	0.014		
	Quantiles						
	0.025	-100.0	10.5	-100.0	22.1		
	0.050	-100.0	27.2	-100.0	36.3		
	0.950	118.5	122.9	486.9	520.4		
	0.975	127.8	131.6	549.9	572.0		
MIRR (%)							
	Mean	-3.0	14.4	17.8	41.2		
	Median	14.5	15.2	38.3	40.5		
	SE	40.9	12.1	51.2	17.9		
	Empirical *p*-value	0.165	0.012	0.167	0.013		
	Quantiles						
	0.025	-100.0	6.9	-100.0	15.6		
	0.050	-100.0	10.1	-100.0	25.1		
	0.950	20.4	20.7	60.3	61.3		
	0.975	21.0	21.2	62.3	63.1		

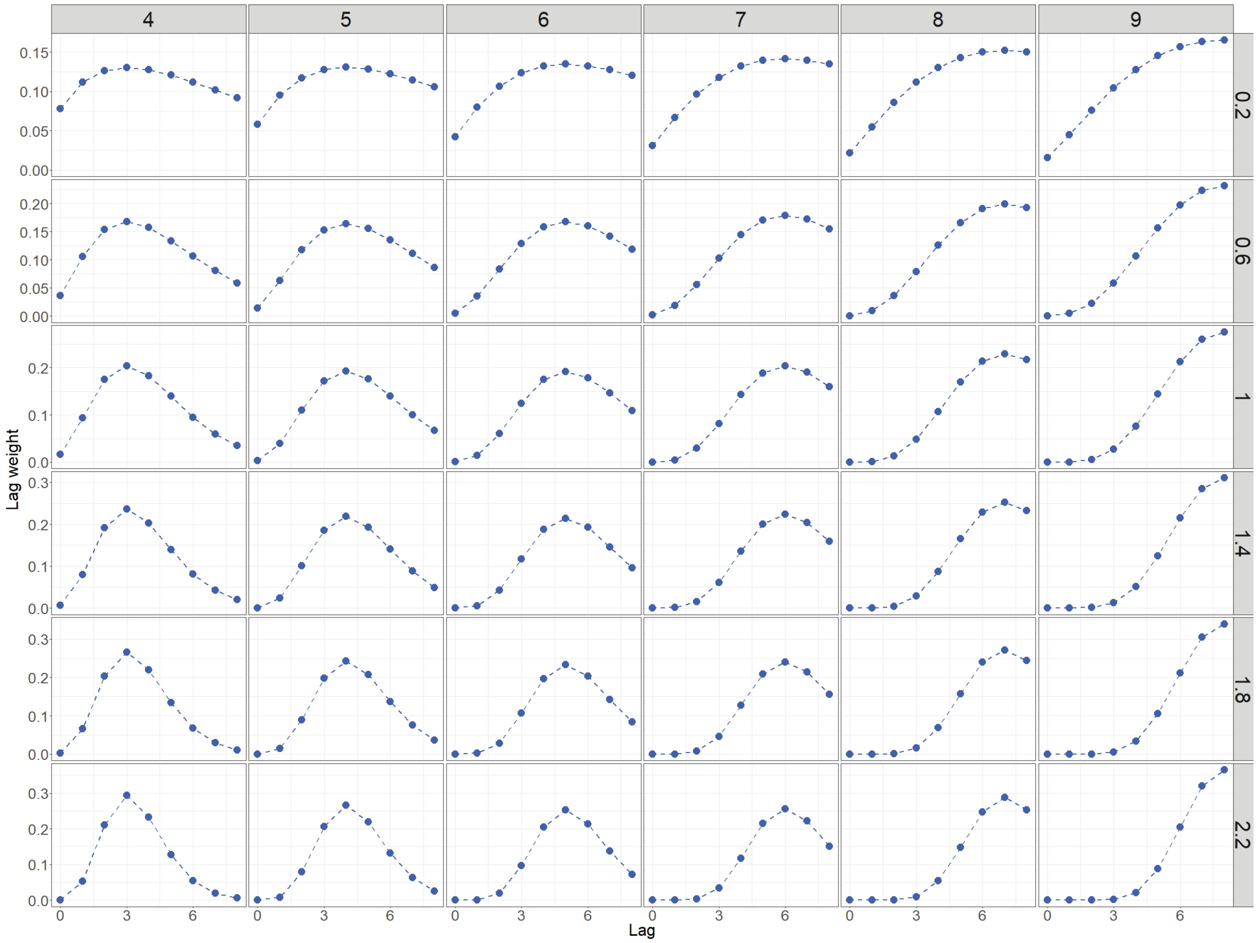
Appendix Figures

App. Figure 2.1. Some research lags used in BMA[48]



[48] Appendix Figures 1 and 2 show a subset of the lags used in BMA. The numbers at the top are peaks of lags, and the numbers on the right are the *b* parameters.

App. Figure 2.2. Some extension lags used in BMA



App. Figure 2.3. Mean and median of parameters as number of draws increases in BHM estimation