The Regional Economics Applications Laboratory (REAL) is a unit in the University of Illinois focusing on the development and use of analytical models for urban and region economic development. The purpose of the Discussion Papers is to circulate intermediate and final results of this research among readers within and outside REAL. The opinions and conclusions expressed in the papers are those of the authors and do not necessarily represent those of the University of Illinois. All requests and comments should be directed to Geoffrey J. D. Hewings, Director, Regional Economics Applications Laboratory, 607 South Mathews, Urbana, IL, 61801-3671, phone (217) 333-4740, FAX (217) 244-9339.

Web page: www.real.illinois.edu/

The Extended Econometric Input-output Model With Heterogeneous Household Demand System

Kijin Kim, Kurt Kratena and Geoffrey J. D. Hewings

REAL 13-T-5 December, 2013

The Extended Econometric Input-output Model With Heterogeneous Household Demand System

Kijin Kim[1,2], Kurt Kratena[2,3], and Geoffrey J. D. Hewings[2]

December 2013

Abstract This paper proposes an extension to the regional econometric input-output model (REIM; Conway, 1990) to which a demand system with age and income parameters is integrated. The extended model addresses concerns about household heterogeneity that has been limited to one representative in the existing REIMs. The initial testing is conducted with a model for the Chicago metropolitan area. First, using aggregate expenditure data by income and age groups, a system of demand equations is constructed: the almost ideal demand system (AIDS) with group fixed effects. Next, the estimated demand systems are linked to the REIM to reflect long-term changes in the age and income distribution of households. The long-range simulation from the extended model takes into account structural changes in expenditure type stemming from changing demographic composition. The extended model further broadens the scope of impact analysis under various scenarios associated with age and income changes.

Key Words: Econometric input-output model, demand system, AIDS, SUR
JEL Classification: C51, C53, D12, J11, R15

1 Introduction

Personal consumer expenditures account for approximately 70 percent of gross domestic product in the US as of 2011, compared to 56 percent on average for OECD countries. Yet most economic models persist in aggregating all the household heterogeneity into one "representative household sector" while, in contrast, industries are often represented by 50 to 500 different sectors. With an ageing population, increasing mobility and widening income inequality becoming critical issues in advanced economies, analysis that highlights their implications for consumer demand is now regarded as a major priority. The regional econometric input-output model (REIM; Conway, 1990; Israilevich *et al.*, 1997) is one of several alternative economic models that provide a way to extensively examine the long-term effects of socio-demographics changes at the regional level. The REIM has its roots in

[1] Corresponding author: kkim96@illinois.edu
[2] Regional Economics Applications Laboratory, University of Illinois, Urbana, IL 61801, USA
[3] Austrian Institute of Economics Research (WIFO), Postfach 91 A1103 Wien, Austria

an empirical macroeconometric model with an integrated input-output component for subnational economies. The combination of dynamic econometric and static input-output approaches offers better forecasting accuracy than the traditional structural econometric models and it also allows inter-industry impact studies with dynamics (Rey, 2000). Based on Conway's methodology (1990), Israilevich *et al.* (1997) further developed the REIM for the Chicago metropolitan area to evaluate the economic impacts with inter-industry spillover reflected through the structure of the input-output table as well as providing an endogenous procedure for updating the input-output structure. One of the caveats in the REIM is that household consumption is limited to a representative consumer mainly due to the absence of detailed consumer expenditures data at the regional level. Thus, the economic effects of changes in household characteristics such as age and income distributions have not been captured so far in the current structure of the REIM.

This paper proposes an extended econometric input-output model for the Chicago region in which an aggregate demand system with parameterized household characteristics is augmented. The integration procedure is as follows: first, using aggregate consumption data from the 1987-2011 annual Consumer Expenditure Survey (CES) and the CPI, we estimate a system of demand equations with age- or income-group fixed effects: the almost ideal demand system (AIDS) of Deaton and Muellbauer (1980a). Income and price elasticities for goods or services are allowed to vary by age or income groups. Next, an integration procedure is proposed by which the demand system is linked to the REIM. In the extended model, distinct spending patterns by cohort[4] are major forces that drive differentiated changes in output, employment and income. Simulation reveals that demographic change (e.g. an ageing population) results in compositional changes in consumption in the long run.

To the best of our knowledge, this is the first attempt to fully integrate the REIM and a demand system that allows heterogeneity in household consumption. Mongelli *et al.* (2010) discuss integration of the AIDS within the static input-output framework. Although Yoon and Hewings (2006) attempted to incorporate the results separately obtained from the REIM and a demand system, this paper provides superior results in that: 1) a generalized

[4] A cohort generally means a group of individuals with *time-invariant* characteristics (e.g. birth cohort; woman born in 1970). However, this study defines a cohort as a group of households with *common* characteristic and "a group" is used interchangeably.

approach to endogenizing a demand system within the REIM is provided; 2) the demand systems are constructed so that they are not only consistent with aggregate demand theory but also parsimonious in terms of empirical estimation. The proposed approach will benefit researchers and regional policy makers to comprehensively understand the economic impacts of changes in age or income distributions.

This paper proceeds as follows. Section 2 describes the structure of the REIM. Section 3 contains brief introductions to the AIDS. Section 4 presents data and describes how the data sources from the REIM and the demand system are matched through a bridge matrix. Section 5 discusses the theoretical aspects of aggregating individual demands deriving the econometric specifications for market demands. Estimation methods and results follow. Section 6 describes the proposed procedure of integrating the demand system into the REIM. Section 7 includes the simulation results and section 8 concludes the paper.

2 Regional econometric input-output model (REIM)

Since its introduction by Israilevich *et al.* (1997), the REIM for the Chicago metropolitan area (CREIM) has been continually maintained and updated by the Regional Economics Application Laboratory. Focusing on subnational regions, the methodology in the REIM is based on a macroeconometric modeling framework in which a static input-output model and dynamic econometric models are integrated. The CREIM has adopted *the coupling strategy* as a way of integration that "reflect[s] the greatest degree of model closure and extent of interaction between the EC [econometric] and IO [input-output] modules" and this approach "results in the most comprehensive representation of regional system" compared to the alternative methods such as the *embedding* and *linking* strategies (Rey, 1998). The integration offers improved forecasting accuracy and inter-industry analysis with dynamics. Characteristics of the REIM are described in greater detail in Rey (2000) and West (1995). A brief description on estimation and forecasting in the CREIM follows.

An overview of the REIM is presented in Figure 1. Exogenous exports and endogenous final demand lead to changes in output. Constant-price actual output (a vector of sectoral output, o_i's; O) is expressed as a function of constant-price expected output (Z) that contains the deterministic structure of the base-year input-output table:

$$Z = \boldsymbol{A}O + \boldsymbol{B}F$$
$$\log(o_i/z_i) = f_i(\cdot) + \varepsilon_i \text{ or } \log(o_i) = f_i'(\log(\mathrm{z_i})) + \varepsilon_i' \tag{1}$$

where $\boldsymbol{A}$ is a matrix of technical coefficients; $\boldsymbol{B}$ is a coefficient matrix normalized so that each column of final demand component adds up to one; $\boldsymbol{F}$ is a matrix of real final demand components. Since personal consumption expenditure data for the Chicago region are not available, it is assumed that for four expenditure types, i.e., auto and parts, other durables, nondurables and services, consumption equations for Chicago and the US have identical functional forms. Consumption expenditures on a per capita basis for the US are first estimated using personal income as one of the explanatory variables. Then, consumption expenditures for Chicago are generated by inserting local personal income into the estimated equations for the US.

<< Insert figure 1 here >>

The elements in $\boldsymbol{A}$ and $\boldsymbol{B}$ are constant since they are based on the base-year input-output table. The stochastic relationship between actual and expected output is one of various ways to overcome the often-criticized constancy of technical coefficients in the input-output approach;[5] the movement of differences between O and Z represents overall changes in technical coefficients over time while O and Z are identical in the base year by construction. Labor productivity defined by output per worker is estimated in the following form:

$$\log(o_i/n_i) = g_i(\cdot) + u_i. \tag{2}$$

Similarly, per capita real income is estimated as

$$\log(y_i/n_i) = h_i(\cdot) + \nu_i. \tag{3}$$

Total population is endogenously determined by labor demand. Five out of six age groups are assumed to follow the national trend and the remaining age 25-44 group is expressed as the residual. Population and income (Y) determine final demand in turn, completing the feedback loop starting from final demand to output, employment, population, income, and again to final demand. To generate forecasts, all of estimated equations are numerically solved for endogenous variables using the Gauss-Seidel algorithm.[6] Long-term forecasts of

[5] See Klein *et al.* (1999, pp. 35-39) for other ways to estimate changes in the IO coefficients over time.
[6] See Klein *et al.* (1999, chapter 5) for the Gauss-Seidel algorithms for nonlinear equations.

exogenous variables (e.g. variables at the national level) were provided by the IHS Global Insight.

3 Almost ideal demand system (AIDS)

The AIDS model of Deaton and Muellbauer (1980) gained popularity from its functional form that allows flexibility in income elasticity as well as substitutability and complementarity among goods. Moreover, it is straightforward to empirically test properties of demands such as homogeneity and symmetry. Derivation of the AIDS model begins from the PIGLOG[7] cost function for a representative consumer, with $c(\boldsymbol{p}, u)$ for utility u and a price vector $\boldsymbol{p}$:

$$\log c(\boldsymbol{p}, u) = \log a(\boldsymbol{p}) + u\, b(\boldsymbol{p}) \tag{4}$$

The translog form of $\log a(\boldsymbol{p})$ and the Cobb-Douglas form of $b(\boldsymbol{p})$ are chosen for a flexible cost function:

$$\begin{gathered} \log a(\boldsymbol{p}) = \alpha_0 + \textstyle\sum_k \alpha_k \log p_k + \frac{1}{2} \sum_k \sum_j \gamma_{kj}^* \log p_k \log p_j \\ b(\boldsymbol{p}) = \textstyle\prod_k p_k^{\beta_k} \end{gathered} \tag{5}$$

Since the cost function is homogeneous of degree one in $\boldsymbol{p}$ and increasing in u, it is required that $\sum_{k=1}^{I} \alpha_k = 1$ and $\sum_k \gamma_{kj}^* = \sum_j \gamma_{jk}^* = \sum_k \beta_k = 0$. Solving (4) for u yields an indirect utility function $v(\boldsymbol{p}, x)$ for total expenditure x as:

$$v(\boldsymbol{p}, x) = \frac{\log x - \log a(\boldsymbol{p})}{b(\boldsymbol{p})} \tag{6}$$

where $x = c(\boldsymbol{p}, u)$ for a utility-maximizing consumer.

By applying Roy's identity and substituting (5) into (6), the AIDS model is given by

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{x}{P}\right) \tag{7}$$

[7] The price independent generalized logarithm (PIGLOG), termed by Muellbaur (1975, 1976), is based on the following results: by assuming that individuals have the PIGLOG cost functions, 1) the resulting *representative* total expenditure depends only on the distribution of expenditures for individuals, not the prices, and 2) it is a more generalized form than the Gorman polar form. The PIGLOG cost function allows exact aggregation of budget shares over individual consumers and nonlinear Engel curves (Deaton and Muellbaur, 1980b; chapter 5 and 6).

where

$$\log P = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2}\sum_k \sum_j \gamma_{kj} \log p_k \log p_j;\ \gamma_{ij} = \frac{1}{2}\left(\gamma_{ij}^* + \gamma_{ji}^*\right). \tag{8}$$

This form is an extension to the Working-Leser model (Leser, 1963; Working, 1943) which takes into account the relationships between the share value and log of total expenditure. If the price index P is proportional to a known price index such as the Stone's price index P^*, i.e., $P^* \equiv \prod_k p_k^{w_k} \approx \phi\, P$ for a constant ϕ, (7) is expressed linearly in parameters, which facilitates a straightforward econometric estimation. Hence, the linear approximate AIDS (LA/AIDS) is defined as

$$w_i = \alpha_i^* + \sum_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{x}{P^*}\right) \tag{9}$$

where $\log P^* = \sum_k w_k \log p_k$ and $\alpha_i^* = \alpha_i + \beta_i \log \phi$

The AIDS and the LA/AIDS satisfy properties of demand functions (Deaton and Muellbauer, 1980b) providing:

$$\begin{array}{rc} \text{Adding up:} & \sum_{i=1}^{I} \alpha_i = 1, \sum_i \gamma_{ij} = 0, \sum_i \beta_i = 0 \\ \text{Homogeneity:} & \sum_j \gamma_{ij} = 0 \\ \text{Symmetry:} & \gamma_{ij} = \gamma_{ji} \end{array} \tag{10}$$

Negative semi-definiteness of the Slutsky matrix cannot be achieved by any restriction on the parameters. However, negativity as well as homogeneity and symmetry are empirically testable while additivity is automatically satisfied in the case of equation-by-equation estimation using the OLS[8].

4 Data[9]

4.1 Consumer Expenditure Surveys

Aggregate household expenditures in the Chicago region[10] are obtained from the 1987-2011

---

[8] For proof, see exercise 1.12 in Deaton and Muellbauer (1980b).

[9] Major features of the Consumer Expenditure Survey described in this section draws partly on the BLS Handbook of Methods (1997, chapter 16).

[10] The Chicago region in the CES covers 14 counties: Cook, DeKalb, Du Page, Grundy, Kane, Kankakee, Kendall, Lake, McHenry, Will (IL); Lake, Porter, Newton (IN); Kenosha (WI). Meanwhile, the CREIM defines the Chicago region as 7 counties in Illinois: Cook, Du Page, Kane, Kendall, Lake, McHenry and Will.

Consumer Expenditure Surveys (CES) by the Bureau of Labor Statistics (BLS). The CES defines consumer units as households representing the US civilian noninstitutional population. Nearly 80 percent of 7,000 households remain in the sample for five successive quarters and then are replaced with new households after the fifth interview (a rotating panel). Each household is randomly drawn to represent 10,000 households in the US. The resulting expenditure data are used to compute the weights in the Consumer Price Index (CPI).

Seven broadly defined categories are used for demand analysis: (1) food and beverages, (2) nondurables and services for housing, (3) durables for housing, (4) durables for transportation, (5) nondurables and services for transportation, (6) health care, and (7) miscellaneous goods and services. A detailed list of goods and services covered in the CES is provided in Table 1. The national CES contains average annual expenditures by income and age groups: quintiles of income (lowest 20 percent to highest 20 percent) and seven age groups (under 25, 25-34, 35-44, 45-54, 55-64, 65-74, over 75). The BLS releases only average expenditure of all consumer units in the Chicago region, thus expenditures by age or income groups require estimation on the basis of available national data: first, by assuming that the shape of the joint distributions for age (or income) and total expenditure in the US and Chicago are identical, it is possible to generate total expenditures for each income and age cohort in Chicago. Next, it is assumed that consumption patterns (i.e., budget shares) in the US and Chicago within the same age (income) cohort are identical.

<< Insert table 1 here >>

Since expenditure data in the CES exist only in dollar amounts (i.e., quantity times unit price), additional price measures are necessary for demand analysis. Price data are obtained from annual CPI for all urban consumers (CPI-U) in the Chicago-Gary-Kenosha area. As shown in Table 1, the categories in the CPI are matched as closely as possible with the CES in the most detailed level of classification, and then are aggregated to higher levels using annual expenditures as weights. In case of the items where the CPIs for the Chicago area are not available, the corresponding indices for the US are inserted instead. The CPIs for education and recreation are available since 1992 and the CPI for vehicle purchases is available since 1998 while the Personal Consumption Expenditures (PCE) prices in the US national accounts for these items are available since 1987. We estimated an ARIMA model for each CPI with the corresponding PCE price as an explanatory variable and used the model to back-calculate earlier prices.

The 2011 total expenditures by income and age groups are shown in Figure 2. Total expenditures across age group show a hump-shaped curve to peak at the 45-54 age groups. Obviously, total expenditure increases as income increases, but with a large jump between the richest and the second richest group. Average budget shares by age and income groups over the period 1984-2011 are compared in Figure 3 and Figure 4. Families with older reference persons (heads of families) tend to allocate more budget relatevely to health care and other goods and services, less to apparel, transportation, and entertainment. Low income families tend to spend relatevely more on housing (mostly rent) and foods. Budget allocation to entertainment and personal insurance and pension rise as family income increase. These findings suggest that it is essential for consumption analysis to take into account heterogeneity of households in each group.

<< Insert figure 2, 3 & 4 here >>

4.2 Classification match between the CES and the CREIM

Private consumption in the CREIM is classified into 47 aggregate types of products (see Table 2), which is based on the categories of the 2009 input-output table for the Chicago region. The 2009 input-output table for the Chicago region was purchased from IMPLAN Group (formerly MIG Inc.). The original IO table is based on the 6-digit North American Industry Classification System (NAICS).[11]

<< Insert table 2 here >>

On the contrary, the CES reclassified for demand system has 7 types of consumer expenditures aggregated from 21 categories. Since estimated demand systems using the data from the CES are to be incorporated in the CREIM, it requires a brige matrix linking the CES and the CREIM. Before considering direct conversion from the CES to the CREIM, it is worth noting that the PCE in the National Income and Product Accounts (NIPA) are compiled separately by two standards: by type of products (NIPA table 2.4.5) and by function (NIPA table 2.5.5). If a bridge matrix connecting the two criteria is available, it would be possible to relate consumer expenditures in purchasers' prices (by function) to production in producers' prices (by type of products). For example, consumers' new car

[11] See MIG, Inc. (2002) for more details on the construction of IO tables by IMPLAN.

purchases are translated by the bridge matrix into car manufacturing, wholesale and retail trade (trade margin), truck, air or rail transportation (transportation margin). Note that expenditures in the CES are recorded from the consumers' viewpoint while those in the CREIM are from the viewpoint of suppliers. Hence, the PCE bridge matrix is used as an intermediate link between the classifications in the CES and the CREIM. The 110 x 83 US PCE bridge matrix for 2010, that relates 110 products to 83 consumption types, was provided by the INFORUM (Interindustry Forecasting project at the University of Maryland).

Matching between the CREIM and the CES proceeds as follows. First, the PCE by function is matched with the CES category. Similarly the PCE by type of products is matched with the CREIM classification. Next, the 110×83 PCE bridge matrix[12] is reduced to 47×7 to be used for linking the classifications between the CREIM and the CES. Finally, a coefficient matrix is generated by dividing each element by its column sum so that the $(i,j)^{th}$ element represents the fraction of a dollar demanded for product i in the CREIM when one dollar is spent on good j in the CES. By assuming the constancy of the coefficient matrix, one can convert 7 expenditure types in the CES to 47 sectors in the CREIM during the whole sample period and the forecast period.

[12] Since we put focus on a region smaller than a country, the following rows and columns in the PCE bridge matrix are discarded: (row) noncomparable imports/scrap, used and secondhand/rest of the world adjustment to final uses; (column) Americans' travel abroad/foreigners' spending in the US/final consumption expenditures of nonprofits.

5 Estimation of demand system

5.1 Aggregation over groups

Aggregate demand equations require additional parameters for income distribution and average household characteristics. The two parameters should be estimated unless cross-sectional microdata are available (Denton and Mountain, 2011). Following Deaton and Muellbauer (1980a), the LA/AIDS in (9) can be rewritten for an individual household with an identifier h as[13]

$$w_{ih} = \alpha_i + \sum\nolimits_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{x_h/k_h}{P}\right) \tag{11}$$

where $h = 1, \ldots, H$ and $\log P$ is the Stone's price index.

The parameter k_h represents a measure of effective household size such as the number of family members and demographic characteristics of family. With the presence of k_h, it is possible to take into account adjusted total expenditure for per-capita level that is specified. Denote average aggregate budget share of good i for cohort c (e.g. household heads in their 30's or households whose income levels are in the lowest 20 percent) by:

$$W_i^c \equiv \frac{\sum_{h\in c} p_i q_{ih}}{\sum_{h\in c} x_h} = \frac{\sum_{h\in c} x_h w_{ih}}{\sum_{h\in c} x_h} = \frac{\sum_{h\in c} x_h w_{ih}}{X^c}$$

for $c = 1, \ldots, C\ (< H)$ and $X^c = \sum_{h\in c} x_h$. Equation (11) can be rewritten as:

$$w_{ih} = \alpha_i + \sum\nolimits_j \gamma_{ij} \log p_j + \beta_i \left[\log\left(\frac{x_h/k_h}{\bar{x}^c}\right) + \log\left(\frac{\bar{x}^c}{P}\right)\right] \tag{12}$$

where $\bar{x}^c$ is average total expenditure for cohort c.

Taking the weighted average of (12) over cohorts with total expenditure as weights yield aggregate demand for cohort c:

$$W_i^c = \alpha_i + \sum\nolimits_j \gamma_{ij} \log p_j + \beta_i \log\left(\frac{\bar{x}^c}{P}\right) + \theta_i^c \tag{13}$$

where

[13] The superscript stars (*) in the constant term and the price index previously appeared in (9) are omitted for convenience.

$$\theta_i^c = \beta_i \left[\sum\nolimits_{h \in c} \frac{x_h}{X^c} \log\left(\frac{x_h}{\bar{x}^c}\right) - \log k^c \right] \qquad (14)$$

for $k^c \equiv \left(\prod_{h \in c} k_h^{x_h}\right)^{1/X^c}$.

The parameter θ_i^c contains not only an income inequality measure but also average household characteristics in cohort *c*. The first term in the parenthesis in (14) represents Theil's income inequality measure for cohort *c*, which has a value of zero in the case of perfect income equality. The second term is the logarithm of the weighted geometric mean of family size in cohort *c*. Since average family size is likely to be positively correlated with aggregate total expenditure, estimation of (13) without θ_i^c produces inconsistent estimates. There are two assumptions that account for the cohort effect on expenditure type *i*, θ_i^c, that enable econometric estimation of the parameter with the currently available data.

First, it is assumed that cohort effects do not change over time. The most important changes in cohort characteristic will stem from family size. Figure 5 & Figure 6 show the trends of average numbers of family members in the US: children under 18, persons 65 and over, earners, and vehicles by age and income of household head.[14] Average family size varies among groups and also features a slight variation or very slowly changing trends for the last two and a half decades. This strongly supports the assumption of time-invariant family composition. In this sense, prices and total expenditure being held constant, θ_i^c represents the long-term average of budget share for each cohort's consumption of good *i*.[15]

<< Insert figure 5 & 6 here >>

Additionally, it is assumed that the income inequality measure for each cohort shares a common linear time trend, but has its own intercept.[16] It turns out that adding the time

---

[14] The CES does not release comparable family characteristics by age and income groups for the Chicago region. Thus, we assume that the national family characteristics are good approximates for city-level characteristics.

[15] Noticeable difference in the long-term average consumption patterns among age or income groups is also observed in Figure 3 and Figure 4. With a limited number of observations (only one observation for each period is available for each cohort), it is not possible to estimate time-varying group effects. Instead, we experimented the followings: 1) a model with cohort effects and time fixed effects, 2) a model with simplified time-specific cohort effects, in which each period is assigned one if expansion or zero if recession. None of the models showed improvement in the BIC than the model with constant cohort effects.

[16] Gini coefficient for the US compiled by the Census Bureau shows a rising trend since the mid 1960s. We calculated the Theil index for each cohort by using the micro data for the US. Estimation results from the model with the Theil index did not show much difference compared to the model with common trends.

trend also captures the effects of average household characteristics that show a rising (or declining) trend such as the percentage of household heads with college degrees and the percentage of female household heads. Hence, the adjusted specification for (13) with a time script and a stochastic error is given by[17]

$$W_{it}^{c} = \alpha_i + \delta_i t + \sum_j \gamma_{ij} \log p_{jt} + \beta_i \log\left(\frac{x_t^c}{P_t}\right) + \theta_i^c + \varepsilon_{it}^c \tag{15}$$

where $\sum_i \delta_i = \sum_i \theta_i^c = 0$ for adding up in addition to (10)

for $c = 1, \dots, C;\ i = 1, \dots, I;\ t = 1, \dots, T$.

5.2 Seemingly unrelated regression (SUR)

For estimation of cohort effects θ_i^c, the observations from time series data for all cohorts given good *i* are stacked in the following way:

$$\underset{\mathrm{CT}\times 1}{\boldsymbol{w_i}} = \underset{\mathrm{CT}\times(\mathrm{I}+2)}{\boldsymbol{X}} \underset{(\mathrm{I}+2)\times 1}{\boldsymbol{\delta_i}} + \underset{\mathrm{CT}\times(\mathrm{C}-1)}{\boldsymbol{D}} \underset{(\mathrm{C}-1)\times 1}{\boldsymbol{\theta_i}} + \underset{\mathrm{CT}\times 1}{\boldsymbol{\varepsilon_i}} \tag{16}$$

$$\begin{bmatrix} \boldsymbol{w_i^1} \\ \boldsymbol{w_i^2} \\ \vdots \\ \boldsymbol{w_i^C} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X^1} \\ \boldsymbol{X^2} \\ \vdots \\ \boldsymbol{X^C} \end{bmatrix} \begin{bmatrix} \alpha_i \\ \gamma_{i1} \\ \vdots \\ \gamma_{iI} \\ \beta_i \end{bmatrix} + \begin{bmatrix} \boldsymbol{1_T} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{1_T} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{1_T} \end{bmatrix} \begin{bmatrix} \theta_i^2 \\ \theta_i^3 \\ \vdots \\ \theta_i^C \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon_i^1} \\ \boldsymbol{\varepsilon_i^2} \\ \vdots \\ \boldsymbol{\varepsilon_i^C} \end{bmatrix}$$

where $\boldsymbol{1_T}$ is a T×1 vector of ones; $\boldsymbol{D}$ is a matrix of dummy variables where the first cohort is the base; $\boldsymbol{w_i^c}$ is a T×1 vector of good *i*'s budget shares for cohort *c* during the sample periods; $\boldsymbol{X^c}$ is a T×(I+2) matrix of ones, prices and total expenditures for cohort *c* during the sample periods; $\boldsymbol{\varepsilon_i^c}$ is a T×1 vector of random errors for cohort *c*.

If the disturbances within a cohort are contemporaneously correlated (i.e., $E[\varepsilon_{it}^c \varepsilon_{js}^c] = \sigma_{ij}$, if $\mathrm{t} = \mathrm{s}$; 0 otherwise), the seemingly unrelated regression (SUR; Zellner, 1962) is an appropriate choice of estimation method for a set of demand equations. In the SUR, it is straightforward to impose cross-restrictions such as symmetry. A system of demand equations for all goods and services is given by

[17] Bar notation on total expenditure is dropped for convenience. In practice, each cohort might face different prices of the same goods (e.g. food) due to aggregation: expenditure composition of goods (e.g. meat and vegetable) within a higher-level classification (e.g. food) could vary by cohort. In this case, p_{jt} and P_t in (15) are replaced with p_{jt}^c and P_t^c.

$$\boldsymbol{W} = (\boldsymbol{I_I} \otimes \boldsymbol{X})\boldsymbol{\delta} + (\boldsymbol{I_I} \otimes \boldsymbol{D})\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{17}$$

$$\begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \vdots \\ \boldsymbol{w}_I \end{bmatrix} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{X} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{X} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_1 \\ \boldsymbol{\delta}_2 \\ \vdots \\ \boldsymbol{\delta}_I \end{bmatrix} + \begin{bmatrix} \boldsymbol{D} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_I \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_I \end{bmatrix}$$

where $\boldsymbol{I_I}$ is an identity matrix of order of I; $E(\boldsymbol{\varepsilon}) = \boldsymbol{0}$. The vector of errors in (17) is assumed to have the following variance-covariance matrix:

$$E(\boldsymbol{\varepsilon\varepsilon}') = \boldsymbol{\Omega} = \begin{bmatrix} \sigma_{11}\boldsymbol{I_{CT}} & \sigma_{12}\boldsymbol{I_{CT}} & \cdots & \sigma_{1I}\boldsymbol{I_{CT}} \\ \sigma_{21}\boldsymbol{I_{CT}} & \sigma_{22}\boldsymbol{I_{CT}} & \cdots & \sigma_{2I}\boldsymbol{I_{CT}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{I1}\boldsymbol{I_{CT}} & \sigma_{I2}\boldsymbol{I_{CT}} & \cdots & \sigma_{II}\boldsymbol{I_{CT}} \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1I} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{I1} & \sigma_{I2} & \cdots & \sigma_{II} \end{bmatrix} \otimes \boldsymbol{I_{CT}} = \boldsymbol{\Sigma_I} \otimes \boldsymbol{I_{CT}}$$

When $\boldsymbol{\Omega}$ is unknown, the feasible generalized least squares (FGLS) estimator is given by

$$\boldsymbol{\eta}^{\boldsymbol{FGLS}} = \left(\boldsymbol{M}'\widehat{\boldsymbol{\Omega}}^{-1}\boldsymbol{M}\right)^{-1}\boldsymbol{M}'\widehat{\boldsymbol{\Omega}}^{-1}\boldsymbol{W}$$

where $\boldsymbol{\eta} = [\boldsymbol{\delta} \ \vdots \ \boldsymbol{\theta}]'$; $\boldsymbol{M} = [\boldsymbol{I_I} \otimes \boldsymbol{X} \ \vdots \ \boldsymbol{I_I} \otimes \boldsymbol{D}]$; $\widehat{\boldsymbol{\Omega}}$ is a consistent estimator of the variance-covariance matrix.

When identical explanatory variables are present in each equation, the FGLS estimation of the full system is identical to the OLS estimation of equation by equation (Zellner, 1962). For the AIDS model, one of the equations must be dropped for estimation because the additivity implies the sum of errors across equations to be zero, which creates the singularity problem of covariance matrix of errors.[18] Parameters in the omitted equation are estimated by using the linear relationship among parameters across equations due to imposed additivity and homogeneity. Estimates are invariant to the choice of the omitted equation when the system of equations is estimated by the maximum likelihood method (Barten, 1969) or iterated FGLS[19]. The estimation of the AIDS is similar to that of the LA/AIDS except that they are nonlinear systems of equations. In this study, the AIDS are estimated by iterated FGLS.[20]

[18] By construction, $\boldsymbol{\varepsilon_i}$'s are linearly dependent since $\sum_i \boldsymbol{\varepsilon_i} = \boldsymbol{0}$ or $\boldsymbol{\varepsilon}'\boldsymbol{1} = \boldsymbol{0}$. Singularity of the covariance matrix follows from the fact that $\mathrm{E}(\boldsymbol{\varepsilon\varepsilon}'\boldsymbol{1}) = \boldsymbol{\Omega 1} = \boldsymbol{0}$ (Greene, 2003; chapter 14).

[19] Iterated FGLS is equivalent to maximum likelihood estimation with normal errors assumed (Oberhofer and Kmenta, 1974).

[20] For empirical estimation, the following Stata commands are used: SUREG for the LA/AIDS, and NLSUR for the AIDS. Refer to Poi (2008, 2002) for examples of NLSUR in Stata. See Greene (2003, chapter 14.4) for more details on the iterative FGLS.

5.3 Estimation results

The AIDS estimates for age and income groups with homogeneity and symmetry constraints are reported in Table 3. A priori value was assigned to α_0 following Deaton and Muellbauer (1980a). [21] Dummy variables identifying groups are included in each equation. Group fixed effects are shown to be highly significant (not reported due to limited space) suggesting that heterogeneity among groups are modeled properly through dummy variables. The estimates of the LA/AIDS are for the most part similar to those of the AIDS and thus are not provided here. With a few exceptions, signs and magnitudes of the coefficients for age and income groups show similar patterns. Similarity in parameters between age and income groups is as expected because the same sampling units are grouped by either age or income. Furthermore, the results support to some extent the expectation that the parameters on prices and total expenditure are assumed to be identical across individual households. Trends measuring income inequality in aggregate demand are significant in food, housing and transportation. Almost none of the explanatory variables except for group identifiers seem to influence the budget shares of health care in both types of groups.

<< Insert table 3 here >>

Figure 7 illustrates the estimates of price and total expenditure elasticities by expenditure type. Calculations are based on the formulae in Table 4 (Green and Alston, 1990). Estimated elasticities are in line with our expectations: all of the estimated price elasticities show negative signs and total expenditure elasticities are distributed just below or above one. For the age-group model, food, housing and transportation are classified as necessities while the others are as luxuries. On the other hand, housing and transportation are classified as necessities for the income-group model. Taylor and Houthakker (2010; chapter 7), one of the similar studies on consumer demands, estimated the AIDS models using the 1996 CES microdata and found that food is the least elastic expenditure type with regard to own-price changes, which is consistent with our finding. They also found that when income rises, the expenditures on food, housing and health care do not increase as

[21] α_0 is the minimum cost of living when prices are unitary at the base year, i.e., $\log c(p,u) = \alpha_0$ for $p = 1$ and $u = 0$. It is chosen in prior to be a number just below the lowest value of log(total expenditure) among groups in 2009, which is 10.

much proportionately as the rise in income. Note that except for the group-specific fixed effects, consumption behaviors, i.e., the responsiveness to prices and total expenditure, are assumed to be the same within age- or income-groups, and thus the intra-group differences in the estimated elasticities are attributed to the variations in the budget shares.

<< Insert table 4 & figure 7 here >>

6 Integrating the demand system into the REIM

Integrating the estimated demand system into the REIM requires additional linkages and blocks. The proposed procedure is intended to make full use of the results from the REIM without altering its main structure. Figure 8 shows a schematic diagram of the extended model where additional components can be found at the bottom. Endogenous personal income[22] in the REIM, prices established in the national market, and demographic fixed effects determine the budget shares for each cohort through the pre-estimated demand system for five types of nondurable goods and services. The numbers of households by age or income groups are estimated using the relationship between population and the number of households. Consumption of an expenditure type by a group is calculated simply by multiplying the average consumption level for the group by the corresponding total number of families. Since the resulting estimates from the demand system follow the CES expenditure classification, it is necessary to convert them to the classification compatible with the REIM. A bridge matrix is implemented for the conversion, resulting in 47 sectors of consumption. New estimates of consumption by sector entail re-estimation of actual output equations as well as re-calculation of expected output. The circled numbers in Figure 8 are used as references for further descriptions that follow.

<< Insert figure 8 here >>

① *Linkage between personal income and total expenditure*

For each cohort, a linear Engel curve is estimated on a per-household basis; it expresses cohort-wise real total expenditure as a function of real personal income, which is common

[22] Personal income comprises total earnings by place of work, dividends, interest, and rent, adjustment for residence, personal current transfer receipt less contribution for government social insurance.

for all cohorts, and a lagged dependent variable:

$$\log(\frac{x_t^c}{P_t}) = \xi_0^c + \xi_1^c \log(\frac{y_t}{H_t}) + \xi_2^c \log(\frac{x_{t-1}^c}{P_{t-1}}) + e_{1t}^c$$

where x_t^c is average total expenditure for cohort *c* in current dollars; P_t is a translog price index in the AIDS model; y_t is total personal income in constant dollars; H_t is total number of households; e_{1t}^c is the error term. ξ_1^c and ξ_2^c can be interpreted as propensity to consume and habit formation in consumption by cohort in a rather loose sense because average income is based on the entire group of households, not on a specific group. Estimated equations for total expenditure by age and income groups are presented in Table 5. Personal income and a lagged dependent variable seem to explain total expenditure by group relatively well in that the coefficients of determination range 0.41-0.85 for age cohorts and 0.40-0.71 for income cohorts. LM tests show that the estimated equations for all groups but the lowest 20 percent income group are free of the first-order autocorrelation in the residuals.

<< *Insert table 5 here* >>

② *Demand system block*

Given real total expenditure linked to income via the Engel curve along with prices, the estimated AIDS model for *nondurables and services* determines the budget share of expenditure type *i* for cohort *c* as:

$$W_{it}^c = \alpha_i + \delta_i t + \sum_j \gamma_{ij} \log p_{jt} + \beta_i \log\left(\frac{x_t^c}{P_t}\right) + \theta_i^c$$

The price index for expenditure type *i*, p_{it}, is forecast outside the sample period by a simple ARIMA model with national price forecasts for total expenditure, durables, nondurables, services, or gasoline as explanatory variables.

③ *Linkage between population and the number of households*

Since the levels of consumption in the demand system are on a per-household basis, the equations for the number of households by age and income must be available in order to derive total consumption. For age cohorts, the CREIM has four groups of population (18-24, 25-44, 45-64 and over 65) which are compatible with the seven groups of households (under 25, 25-34, 35-44, 45-54, 55-64, 65-75 and over 75) in the consumption estimation. We expect the ratios of population to the number of households to be stationary, moving within the range of 2 to 5. Seven group-specific equations for the log-ratios are constructed as:

$$\log(\frac{POP_t^{c'}}{H_t^c}) = \pi_0^c + \pi_1^c \log(\frac{POP_{t-1}^{c'}}{H_{t-1}^c}) + e_{2t}^c$$

where $POP_t^{c'}$ is population for cohort c'; H_t^c is the number of households for cohort c; e_{2t}^c is the error term. Each demand equation by income cohort represents consumption patterns of households within an income quintile. Thus, once a single equation for the number of households in total is established, each income group simply has one fifth of total number of households. For income cohorts, we employ a functional form for the log-ratio of total population and total number of households identical to that in age cohort. Table 6 presents estimations results for the log-ratios of population to the number of households. Group-specific equations by age are presented in column (1)-(7) and the ratio of totals is given in column (8). Adjusted R-squares of 0.34-0.80 imply that the proposed AR(1) form adequately captures the short-term movements of the ratios.

<< Insert table 6 here >>

④ *Determination of consumption in the CES*

Real consumption of type *i* is obtained by summing over cohort deflated expenditure on type *i* for all households in cohort *c*:

$$C_{it}^{CES} = \sum\nolimits_c (C_{it}^c / p_{it}) H_t^c$$

where $i = 1, \cdots, 5$; $C_{it}^c = x_t^c W_{it}^c$. Similarly, summation over expenditure type yields real consumption by all households in cohort *c*. Note that all consumption expenditures so far include only nondurables goods and services classified in the CES. Hence, they require conversion to the classification consistent with the REIM, as described in the next step.

⑤ *Bridge matrix: conversion to consumption in the REIM*

Estimates of durables goods in the REIM are preserved due to lack of available data while those of nondurables and services are replaced with the CES data. Conversion to real consumption of sector i in the REIM is accomplished via the bridge matrix:

$$C_{it}^{REIM} = (b_{i1}C_{1t}^{CES} + \cdots + b_{i5}C_{5t}^{CES}) + (b_{i6}D_t^1 + b_{i7}D_t^2)$$

where $i = 1, \cdots, 47$; b_{ij} is the *(i,j)*th element of the 47×7 coefficient matrix described in section 4.2; D_t^1 and D_t^2 are auto and parts and other durables determined within the REIM.

⑥ *Re-estimation of actual and expected outputs*

Expected output, a linear combination of actual output and final demand components, needs to be updated due to newly generated estimates of consumption by sector. Accordingly new equations are estimated relating actual output to expected output.

7 Simulations

7.1 Baseline solutions

The long-range forecasts for the next 30 years or so, 2012-2040, are generated by numerically solving the system of non-linear equations. The data are based on the observations during 1987-2011 in the CES in addition to the 1969-2011 observations for final demand, output, income, employment and population in the existing REIM. The baseline solutions for select variables in the age-group model are presented in Table 8.[23]

<< Insert table 8 here >>

The long-term forecasts of average budget shares by expenditure type by cohort are drawn in Figure 9. Thses *determistic* simulations assume no uncertainties in the system of estimated equations and thus produce only a single solution for each endogenous variable. On the other hand, *stochastic simulations* are useful to capture *ex ante* forecast errors embedded in the disturbances, which can be represented in the form of confidence interval[24]. In practice, uncertainties from the disturbances are measured by adding to the estimated

[23]. Except for consumption shares by income group, the baselines solutions from the income-group model are not presented here due to limited space.

[24] Additional sources of forecast errors include uncertainties in (1) the coefficient estimates and (2) the forecasts of exogenous variable, and (3) the misspecification of the model (Fair, 1980). Obviously, the confidence interval widens as more sources of forecast errors are incorporated in the simulations.

equations random numbers drawn from the Gaussian distribution with the sample mean and variance of the residuals before solving the system (Klein *et a*l., 1999; chapter 5). Figure 10 depicts the forecast levels and growth rates of major endogenous variables illustrated with the 95 percent confidence bands obtained from 500 replications of stochastic simulations.

<< Insert figure 9 & 10 here >>

Real income is expeted to show an annual growth rate of 2 percent on average over the next 30 years. Consumption of nondurables and services show a similar growth path during the same periods because personal income is the major determinant of spending in the demand equations. In the extended REIM, structural changes in consumption patterns stem mainly from changing demographic composition. The bottom right plot in Figure 10 depicts the outlook for the number of household by age of family heads in the Chicago region. As baby boomers age, the number of households with family heads of age 65 and above is expected to grow more rapidly than any other age groups. Elderly households with age 65 and over are forecast to reach 1.5 million households, comprising approximately 30 percent of total households by 2040, compared to 20 percent in 2011. As a result, their contribution to consumption growth is expected to continue to rise as well: the consumption share of elderly families is expected to rise to 23 percent by 2040 from 17 percent in 2011. In contrast, the consumption shares by the 45-64 age group are forecast to decline to 34 percent by 2040 from 43 percent in 2011.

Historically, households with elderly heads have been likely to allocate more budget to housing and health care than other age groups, as already shown in Figure 3. If this is the case over the next years, total expenditures on housing and health care are expected to increasingly take up larger portion of total consumption. The long-term forecast shows that the consumption of housing rises to 39 percent by 2040 from 35 percent in 2011. However, it is not in line with our expectations that consumption share of health care shows a declining trend. It is because total real expenditure on housing increase more rapidly than that on health care even if real consumption of heath care in levels does increase. Additionally, the price of services used to deflate health care spending during the out-of-sample periods are assumed to rise at a faster rate than the deflators for other expenditure types.

Unlike age cohorts, income groups show only the slightest variation over time in consumption shares since each group represents exactly 20 percent of total households at any point in time. Though baseline solutions from the income-group model do not provide much insights, it becomes useful when evaluating changes in income distribution, which is described in the following section.

7.2 Scenario analysis

The Chicago region has the most populous counties in Illinois, accounting for approximately 70 percent of total population in the state, and shows highly active migration flows. According to the recent statistics for migration flows in the U.S. given in Table 7, the Chicago region had a net annual out-migration of 100,546 residents on average during 2006-2010; 175,170 in-migrants and 275,716 out-migrants. The total number of people who moved in or out of the Chicago region in a single year accounts for more than 5 percent of total population in the region. In the extended REIM, various scenarios can be simulated by altering the distributions of age or incomes groups, a procedure that was not possible in the existing REIM due to the assumption of one homogeneous household. The EREIM provides a useful analytical tool to evaluate the effects of migration of households whose main characteristics differ by age or income.

<< *Insert table 7 here* >>

Table 9 presents total impacts on consumption, output, income and employment provided that 1,000 households under the same age or income groups in-migrate to the Chicago region in 2015. Total impacts encompasse direct (or immediate changes due to inflow of housholds), indirect (or supplier-induced), and induced (or income-induced) impacts. Inflows of households initially stimulate local consumption. Then, output rises to meet the increased consumer demand, and employment required for additional production are created. The positive income shock due to job growth in turn induces additional consumption. For example, the increase of the 45-55 age group by 1,000 households induces $83 million of total consumption, $113 million of output, $27 million of income, and 587 jobs in the Chicago region. The increase in the youngest households (under 25) to the same extent makes an impact only half the size of total impacts by the 45-55 age group. As for income group, suppose in-migration of 1,000 households whose level of income corresponds to the highest 20 percent in the income distribution of Chicago residents. The inflow is forecast to lead to $139 million of total consumption, $192 million of output, $47 million of income, and 1,007 jobs. The total impacts due to inflows of the poorest group are less than one fourth of total impacts by the wealthiest group. Annual trends of group-specific total impacts of household inflows on consumption are plotted in Figure 11. It shows a gradual rising trend as the Chicago economy grows. Interestingly, the richer a group is, the more rapidly the impact of inflows on consumption grows since high-income groups generate larger secondary effects. Note that total impacts are initiated by consumption and thus the direct changes in labor supply and income due to inflows are not taken into account here.

<< Insert table 9 & figure 11 here >>

The bottom panel of Table 9 shows total impacts on consumption of nondurables and services by type in 2015 due to the increase of movers into the region. Group-specific expenditure shares by type resulting from the population inflows are provided in the parentheses. Furthermore, the expenditure shares and their confidence intervals are plotted in Figure 12 in order to facilitate statistical comparison among groups. Since each cohort shows a unique spending pattern, inflows of households in different cohorts generate compositional differences in expenditure types. For instance, the differences in each expenditure category between the youngest and the eldest are noticeable even though there does not appear to be a large difference in *total* impacts of inflows between the two groups ($35.8 vs. $38.2 million). Consequently, there would be different outcomes on production and labor demand sector by sector. Especially, the contrast of spending on health care and food is worth noting: health care spending increases by 1.2 million (amounting to 3 percent of total change) due to the youngest in-migrants as opposed to $5.3 million (14 percent) due to the eldest in-migrants. The inflow of the under-25 group is expected to lead to an increase in local food consumption by $6.2 million (17 percent) while the inflow of the over-75 group results in additional spending of $5.5 million (14 percent). For income groups, as the poorest group moves to Chicago, 41 percent of the total impact on consumption is concentrated on housing, compared to 34 percent for the richest. Inflow of the highest income-quintile group stimulates spending on miscellaneous goods and services, accounting for 39 percent of the total change, compared to 25 percent for the lowest income-quintile group.

<< Insert figure 12 here >>

8 Conclusions

Since its application to Washington by Conway (1990), the regional econometric input-output model has been applied to several other subnational regions in the US and has proven its usefulness for forecasting and impact study. Due to lack of regional data, however, analysis on consumption in the REIM has been limited to one representative household. This paper proposes an extended REIM for the Chicago region that integrates the existing REIM and the demand system that allows household heterogeneity by utilizing actual household expenditure survey information. The integration requires estimation of a demand system and a bridge matrix converting the estimated consumption demand to the classification in the REIM. The proposed approach will benefit regional modelers in that integration procedure can be applied without difficulty to any regional econometric model with a similar structure. Furthermore, with the modeled structure of inter-regional spillovers, it is possible to extend its application to multi-regional models.

As shown in the long-range simulation, the extended model properly captures the fact that the changing demographic composition results in structural changes in consumer spending pattern. As population ages, the contribution to consumption growth by elderly households is expected to continue to grow. As a result, the goods and services consumed by the elderly group increase their market size. With the aid of an augmented demand system, the extended REIM enables us the evaluation of the economic impacts of various scenarios associated with demographic changes. For example, experiments on in-migration of households in each cohort show that affected sectors in terms of consumption, production and labor demand vary by cohort characteristics even though the total impacts might not be so different. These types of simulation exercises can help regional policy makers analyze the long-term consequences of regional policies regarding economic development, migration, and income inequality.

Limitations of this study include the imperfect classification match between the CES and the REIM. There does not exist a bridge matrix directly linking the household expenditure survey and the PCE of national accounts due to their underlying methodological differences. The paper attempts to address this issue by using the PCE bridge matrix as the link between two classifications of different kinds. One of the limitations is associated with the small number, only five, of expenditure types in the demand system relative to 47 sectors in the REIM. This might be relaxed by imposing additional restrictions on the structure of complementarity and substitutability (though how to justify the structure would remain a problem) and thus securing more degrees of freedom for reliable estimation.

One of the issues left for future research is to model demand for durables goods. Consumer choice for durable goods requires a different approach. Intertemporal choice plays a more important role for durables than for nondurables and services since the presence of stocks in the previous period affects present consumption of durables. Next, although net migration in the CREIM is treated simply as a residual, i.e., population change less net births, it will require more attention when the model is extended to multiple regions, especially for regions with active inter-regional migration flows like states or counties in the US.

Acknowledgements

We acknowledge the support of the EU center, University of Illinois at Urbana-Champaign in the form of a Graduate Student Research Travel Grant. The PCE bridge matrix for 2010 was kindly provided by Dr. Douglas Meade from the Interindustry Forecasting project at the University of Maryland (INFORUM).

References

Barten A.P. (1969) Maximum Likelihood Estimation of a Complete System of Demand Equations. *European Economic Review, 1*, 7-73.

Bureau of Labor Statistics (1997) BLS Handbook of Methods. US Department of Labor.

Conway, R. (1990) The Washington Projection and Simulation Model: A Regional Interindustry Econometric Model. *International Regional Science Review*, 13, 141-165.

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

Deaton, A. and Muellbauer, J. (1980b) Economics and Consumer Behavior. New York, Cambridge University Press.

Denton, F.C. and Mountain, D.C. (2011) Exploring the Effects of Aggregation Error in the Estimation of Consumer Demand Elasticities. *Economic Modelling*, 28, 1747-1755.

Fair, R.C. (1980) Estimating the Expected Predictive Accuracy of Econometric Models. International Economic Review, 21, 355-378

Green, R. and Alston, J. (1990) Elasticities in AIDS models. *American Journal of Agricultural Economics*, 72, 442-445.

Greene, W.H. (2003) Econometric analysis, 5th. Ed. Upper Saddle River, NJ.

Israilevich, P., Hewings, G., Schindler, G., and Mahidahra, R. (1997) Forecasting Structural Change with a Regional Econometric Input-Output Model. *Journal of Regional Science*, 37, 565-590.

Klein, L.R., Welfe, A., and Welfe, W. (1999) Principles of macroeconometric modeling. Elsevier.

Leser, C.E.V. (1963) Forms of Engel functions. *Econometrica*, 31,694-703.

MIG, Inc., IMPLAN System (software and data), 502 2nd Street, Suite 301, Hudson, WI 54016.

MIG, Inc. (2002), Elements of the Social Accounting Matrix, MIG IMPLAN Technical Report TR-98002.

Mongelli, I., Neuwahl, F., and Reuda-Cantuche, J. (2010) Integrating a Household Demand System in the Input-output Framework. Methodological Aspects and Modeling Implications. *Economic Systems Research*, 22, 201-222.

Muellbauer, J. (1975) Aggregation, Income Distribution and Consumer Demand. *The Review of Economic Studies*, 42, 525-543.

Muellbauer, J. (1976) Community Preferences and the Representative Consumer. *Econometrica*, 44, 979-999.

Oberhofer, W., and Kmenta, J. (1974) A General Procedure for Obtaining Maximum Likelihood Estimates in Generalized Regression Models. *Econometrica*, 42, 579-590.

Poi, B.R. (2002) From the help desk: Demand System estimation. *The Stata Journal*, 2, 403-410.

Poi, B.R. (2008) Demand-system estimation: Update. *The Stata Journal*, 2, 403-410.

Rey, S. (1998) The Performance of Alternative Integration Strategies for Combining Regional Econometric and Input-Output Models. *International Regional Science Review*, 21,1-36.

Rey, S. (2000) Integrated Regional Econometirc+Input-output Modeling: Issues and Opportunities, *Papers in Regional Science*, 79, 271-292.

Stone, R. (1953) The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom, 1920-1938, Vol. 1, Cambridge.

Stone, R. (1954) Linear Expenditure Systems and Demand Analysis: an Application to the Pattern of British Demand. *The Economic Journal*, 64, 511-527.

Taylor, L.D., and Houthakker, H.S. (2010) Consumer Demand in the United States: Prices, Income, and Consumption Behavior. Springer.

West, G. (1995) Comparison of Input-output, Input-output+Econometric and Computable General Equilibrium Impact Models at the Regional Level. *Economic Systems Research*, 7, 209-227.

Working, H. (1943) Statistical Laws of Family Expenditure. *Journal of the American Statistical Association*, 38, 43-56.

Yoon, S.G. and Hewings, G. (2006) Impacts of Demographic Changes in the Chicago Region. REAL Discussion Paper 06-T-7.

Zellner, A. (1962) An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias, *Journal of the American Statistical Association*, 57, 348-368.

Tables

Table 1. Classifications in the CES and CPI

Consumer Expenditure Survey			Consumer Price Index	
New Category (7)	2011 Share (%)	Description (share, %)	Description	Geo-coverage
Food and beverages	13.2	Food (94.2)	Food	Chicago
		Alcoholic beverages (5.8)	Alcoholic beverages	Chicago
Housing (ND+S)	33.1	Shelter (68.6)	Shelter	Chicago
		Utilities, fuels, and public services (20.9)	Fuels and utilities	Chicago
		Household operations (7.0)	Housing	Chicago
		Housekeeping supplies (3.5)		
Housing (D)	2.6	Household furnishings and equipment (100)	Household furnishings and operations	Chicago
Transportation (D)	4.6	Vehicle purchases (100)	New and used motor vehicles	US city avg
Transportation (ND+S)	9.9	Gasoline and motor oil (42.8)	Motor fuel	Chicago
		Other vehicle expenses (42.2)	Transportation	Chicago
		Public transportation (15.1)	Public transportation	US city avg
Health care	7.1	Health care (100)	Medical care	Chicago
Miscellaneous	29.6	Apparel (12.1)	Apparel	Chicago
		Entertainment (18.2)	Recreation	Chicago
		Reading (0.7)		
		Education (10.4)	Education	US city avg
		Personal insurance and pension (38.8)	All items	Chicago
		Personal care (4.3)	Personal care	US city avg
		Tobacco products (1.5)	Tobacco and smoking products	US city avg
		Miscellaneous (4.7)	Miscellaneous personal services	US city avg
		Cash contribution (9.4)	All items	Chicago

Notes: "D", "ND" and "S" are short for durables, nondurables and services, respectively.

Table 2. Classifications in the CREIM

No.	Type of Product	2009 Consumer exp.($Mil)
1	Livestock and Other Agricultural Products	74
2	Agriculture, Forestry and Fisheries	27
3	Mining	162
4	Utilities	4,042
5	Construction	-
6	Food and Kindred Products	5,323
7	Tobacco Product Manufacturing	2,203
8	Apparel and Textile Products	199
9	Leather and Leather Products	9
10	Lumber and Wood Products	44
11	Paper and Allied Products	168
12	Printing and Publishing	1,234
13	Petroleum and Coal Products	3,254
14	Chemicals and Allied Products	4,083
15	Rubber and Misc. Plastics Products	220
16	Stone, Clay, and Glass Products	51
17	Primary Metals Industries	4
18	Fabricated Metal Products	58
19	Industrial Machinery and Equipment	33
20	Computer and other Electric product, component manufacturing	343
21	Transportation Equipment Manufacturing	253
22	Furniture and Related Product Manufacturing	206
23	Miscellaneous Manufacturing	452
24	Wholesale Trade	10,970
25	Retail Trade	25,355
26	Air Transportation	1,803
27	Railroad Transportation and Transportation Services	666
28	Water Transportation	285
29	Truck Transportation and Warehousing (+Waste and remediation services)	2,131
30	Transit and Ground Passenger Transportation	888
31	Pipeline Transportation	30
32	Information (except 33 sector)	3,873
33	Motion Picture and Sound Recording Industries	813
34	Finance and Insurance	25,663
35	Real Estate	49,216
36	Professional and Management services and other support services	7,284
37	Educational Services	9,298
38	Health Care	45,867
39	Social Assistance	4,189
40	Arts, Entertainment, and Recreation	3,955
41	Accommodation Services	163
42	Food Services	14,512
43	Repair and Maintenance	2,538
44	Personal and Laundry Services	4,339
45	Membership Organizations and Private Households	7,005
46	FGE (Federal government)	46
47	SLGE (State and local governments)	2,950
TOTAL		**246,285**

Table 3. Estimated AIDS models with homogeneity and symmetry constraints

	Food	Housing	Transportation	Health care	Misc.
			(Model with Age-group dummy variables)		
Price(food)	**0.099** (0.031)**	-0.016 (0.017)	-0.004 (0.006)	-0.020 (0.019)	-0.059* (0.028)
Price(housing)	-0.016 (0.017)	**0.034* (0.017)**	-0.005 (0.006)	-0.009 (0.013)	-0.005 (0.020)
Price(trans.)	-0.004 (0.006)	-0.005 (0.006)	**0.036** (0.005)**	-0.008 (0.005)	-0.020* (0.009)
Price(health.)	-0.020 (0.019)	-0.009 (0.013)	-0.008 (0.005)	**0.013 (0.023)**	0.024 (0.029)
Price(misc.)	-0.059* (0.028)	-0.005 (0.020)	-0.020* (0.009)	0.024 (0.029)	**0.061 (0.048)**
Real tot. exp.	**-0.022** (0.007)**	**-0.044** (0.010)**	**-0.002 (0.005)**	**0.011 (0.007)**	**0.057** (0.012)**
Constant	0.216** (0.012)	0.306** (0.009)	0.132** (0.004)	0.023 (0.014)	0.323** (0.020)
Trend	-0.001* (0.001)	0.002** (0.000)	0.000 (0.000)	0.000 (0.001)	-0.001 (0.001)
			(Model with income-group dummy variables)		
Price(food)	**0.085** (0.032)**	-0.040* (0.018)	-0.017** (0.006)	-0.060** (0.020)	0.032 (0.030)
Price(housing)	-0.040* (0.018)	**0.043* (0.022)**	0.017* (0.007)	0.008 (0.015)	-0.028 (0.023)
Price(trans.)	-0.017** (0.006)	0.017* (0.007)	**0.051** (0.004)**	-0.008 (0.005)	-0.044** (0.008)
Price(health.)	-0.060** (0.020)	0.008 (0.015)	-0.008 (0.005)	**0.040 (0.025)**	0.020 (0.032)
Price(misc.)	0.032 (0.030)	-0.028 (0.023)	-0.044** (0.008)	0.020 (0.032)	**0.020 (0.050)**
Real tot. exp.	**0.042** (0.012)**	**-0.027 (0.017)**	**-0.023** (0.008)**	**0.003 (0.011)**	**0.005 (0.017)**
Constant	0.183** (0.013)	0.360** (0.010)	0.111** (0.004)	0.089** (0.016)	0.256** (0.022)
Trend	0.000 (0.001)	0.002** (0.000)	0.000* (0.000)	0.000 (0.001)	-0.001 (0.001)

Note: 1) Standard errors are in parentheses; 2) Prices and real total expenditures are in logarithms; 3) Each equation includes dummy variables for age or income groups; 4) *p<.05; ** p<.01

Table 4. Formulae for elasticities

Functional form	Total expenditure elasticity (e_i^c)	Uncompensated Price Elasticity (η_{ij}^c)
LA/AIDS	$1+\frac{\beta_i}{w_i^c}$	$-\delta_{ij}+\frac{\gamma_{ij}}{w_i^c}+(1-e_i^c)\left(w_j^c+\sum_k w_k^c \log\bar{p}_k\left(\eta_{kj}^c+\delta_{kj}\right)\right)$
AIDS	$1+\frac{\beta_i}{w_i^c}$	$-\delta_{ij}+\frac{\gamma_{ij}}{w_i^c}+(1-e_i^c)\left(\alpha_j+\sum_k \gamma_{kj}\log\bar{p}_k\right)$

Notes: δ_{ij} is the Kronecker delta ($\delta_{ij}=1$ for $i=j$; 0 otherwise); Compensated price elasticity is written as $\eta_{ij}^{c*}=\eta_{ij}^c+e_i^c w_j^c$

Table 5. Estimated equations for total expenditures[1)] by group

Group	Income[2)]		First-order lag		Constant		Adj. R sq.	LM F-stat[3)]
(Age group)								
Under 25	0.274**	(0.10)	0.587**	(0.14)	-3.043*	(1.08)	0.698	0.003
25-34	0.207**	(0.07)	0.566**	(0.13)	-2.050**	(0.72)	0.737	0.062
35-44	0.139	(0.07)	0.712**	(0.13)	-1.315	(0.75)	0.644	0.236
45-54	0.086	(0.06)	0.604**	(0.16)	-0.587	(0.68)	0.410	0.156
55-64	0.204**	(0.07)	0.669**	(0.11)	-2.058**	(0.73)	0.815	1.224
65-75	0.231*	(0.09)	0.731**	(0.11)	-2.490*	(0.95)	0.854	0.038
Over 75	0.216*	(0.10)	0.666**	(0.12)	-2.368*	(1.09)	0.796	0.222
(Income group)								
Lowest 20%	0.135*	(0.06)	0.417*	(0.17)	-1.538*	(0.72)	0.401	6.347*
Second 20%	0.172*	(0.07)	0.376*	(0.18)	-1.742*	(0.75)	0.474	0.609
Third 20%	0.143*	(0.06)	0.492**	(0.16)	-1.300	(0.63)	0.519	0.416
Fourth 20%	0.119*	(0.05)	0.538**	(0.16)	-0.918	(0.56)	0.536	3.741
Highest 20%	0.173**	(0.06)	0.543**	(0.14)	-1.312*	(0.56)	0.714	0.060

Notes: 1) log(total expenditure/price index); 2) log(total personal income/total number of households); 3) Breusch-Godfrey's LM test for first-order autocorrelation; 4) Standard errors in parentheses; 5) Sample periods: 1987-2011; 6) * p<.05; ** p<.01

Table 6. Estimated equations for log-ratios of population to the number of households

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Population	18-24	25-44	25-44	45-64	45-64	>65	>65	Total[2)]
#HHs	<25	25-34	35-44	45-54	55-64	65-75	>75	Total
1st-order lag	0.825**	0.834**	0.741**	0.858**	0.794**	0.863**	0.577**	0.744**
	(0.14)	(0.09)	(0.14)	(0.10)	(0.11)	(0.09)	(0.17)	(0.12)
Constant	0.217	0.249	0.343	0.151	0.286	0.135	0.462*	0.233*
	(0.16)	(0.13)	(0.18)	(0.11)	(0.14)	(0.09)	(0.19)	(0.11)
Adj. R sq.	0.613	0.785	0.552	0.755	0.711	0.795	0.342	0.701
LM F-stat[1)]	0.713	1.364	2.810	4.586*	5.765*	0.883	4.272	4.076

Notes: 1) Breusch-Godfrey's LM test for first-order autocorrelation; 2) Year dummy variables for 2000 and 2006 are included in the equation; 3) Standard errors in parentheses; 5) Sample periods: 1987-2011; 6) * p<.05; ** p<.01

Table 7. Annual migration flows in the US

(Unit: 1,000 person)

Origin \ Dest.	Chicago-IL	Others-IL	ROUS	Total out.	Total net
Chicago-IL*	0	66.0	209.7	275.7	-100.5
Others-IL	28.6	0	93.9	122.4	25.1
ROUS	146.6	81.6	0	228.2	75.4
Total In.	175.2	147.5	303.6	626.3	0.0
Population	8,376	4,243	296,129	308,748	-

* The Chicago region includes 7 counties in Illinois: Cook, Du Page, Kane, Kendall, Lake, McHenry and Will; Source: 2006-2010 5-year American Community Survey (ACS)

Table 8. Baseline solutions for select endogenous variables

Variables	Units*	Observed		Forecast		
		1990-99	2000-11	2012-19	2020-29	2030-40
Output	$2009 Bil.	684.9	907.2	1,163.3	1,594.3	2,263.5
	(growth, %)	(2.3)	(2.6)	(3.2)	(3.2)	(3.2)
Income	$2009 Bil.	240.7	266.8	315.7	386.0	480.3
	(growth, %)	(2.3)	(0.9)	(2.1)	(2.0)	(2.0)
Employment	1,000 per.	4,690	4,773	5,456	6,385	7,592
	(growth, %)	(1.5)	(0.2)	(1.7)	(1.6)	(1.6)
GDP	$2009 Bil.	384.0	425.3	532.4	695.3	872.5
	(growth, %)	(2.5)	(0.9)	(2.8)	(2.7)	(2.0)
Consumption	$2009 Bil.	166.5	235.5	286.1	369.0	474.3
	(growth, %)	(0.1)	(3.2)	(2.5)	(2.6)	(2.3)
Nondur. & Serv.	$2009 Bil.	130.5	172.7	204.1	249.7	299.2
	(growth, %)	(-0.7)	(2.6)	(2.1)	(2.0)	(1.6)
(by type)						
Food	share (%)	16.5	15.6	14.2	12.7	11.2
Housing	share (%)	33.7	34.8	36.3	37.7	39.0
Transportation	share (%)	12.3	9.8	8.9	8.8	8.4
Health care	share (%)	7.2	7.1	6.3	5.9	5.3
Misc.	share (%)	30.2	32.7	34.3	34.9	36.1
(by age)						
Under 25	share (%)	4.2	3.7	3.9	4.2	4.5
25-34	share (%)	17.1	16.0	15.3	15.3	15.5
35-44	share (%)	26.0	20.4	22.4	22.8	23.1
45-54	share (%)	24.3	23.9	24.4	21.9	20.5
55-64	share (%)	13.4	19.1	15.8	14.2	13.7
65-75	share (%)	8.7	10.4	11.0	13.2	14.0
Over 75	share (%)	6.3	6.5	7.2	8.5	8.7
(by income)						
Lowest 20%	share (%)	9.0	9.0	8.7	8.6	8.5
Second 20%	share (%)	12.8	13.0	12.5	12.4	12.3
Third 20%	share (%)	17.0	17.0	16.9	16.7	16.5
Fourth 20%	share (%)	23.3	22.9	22.9	22.8	22.7
Highest 20%	share (%)	38.0	38.0	39.0	39.5	40.0
(Exogenous var.)						
PCE Prices	growth, %	2.4	2.2	2.4	2.1	2.3
Durables	growth, %	-0.2	-2.2	-1.2	-1.9	-2.0
Nondurables	growth, %	1.9	2.3	2.3	1.6	1.8
Gasoline	growth, %	1.3	6.7	5.3	-0.1	1.3
Services	growth, %	3.3	3.0	2.4	2.5	2.4

* Levels and shares are for the last period of the sample years. Growth rates are averages during the periods. All figures but shares by income group are obtained from the EREIM for age groups. The differences in forecasts between the income-group EREIM and the age group EREIM are trivial.

Table 9. Total impacts of an 1,000 household increase by group on select variables

(Impacts on select variables)

(Unit: $2009 Mil., 1,000 person)

(2015)	Consumption	Output	Income	Employment
		(Age group)		
Under 25	38.9	53.8	12.8	0.281
25-34	64.3	87.2	20.6	0.449
35-44	80.0	108.8	25.9	0.563
45-54	82.6	112.6	27.2	0.587
55-64	70.6	95.7	23.1	0.499
65-75	56.2	75.3	18.1	0.390
Over 75	41.2	53.9	12.8	0.273
		(Income group)		
Lowest 20%	31.2	42.6	9.9	0.218
Second 20%	44.7	61.0	14.4	0.314
Third 20%	60.7	83.3	19.8	0.433
Fourth 20%	82.2	113.2	27.2	0.593
Highest 20%	139.2	191.8	46.5	1.007

(Impacts on consumption by types - nondurables and services)

(Unit: $2009 Mil., %)

(2015)	Food	Housing	Trans.	Health.	Misc.	Total
			(Age group)			
Under 25	6.2	13.0	3.5	1.2	12.0	35.8
	(17.2)	(36.3)	(9.6)	(3.4)	(33.4)	(100.0)
25-34	9.1	22.9	5.5	2.5	19.3	59.3
	(15.3)	(38.6)	(9.3)	(4.2)	(32.6)	(100.0)
35-44	11.1	27.3	6.6	3.5	25.3	73.8
	(15.1)	(37.0)	(8.9)	(4.7)	(34.3)	(100.0)
45-54	11.0	26.2	7.1	4.0	27.7	76.1
	(14.5)	(34.5)	(9.4)	(5.2)	(36.4)	(100.0)
55-64	9.4	22.2	6.1	4.5	22.8	65.0
	(14.5)	(34.1)	(9.4)	(6.9)	(35.1)	(100.0)
65-75	7.8	18.2	4.8	5.5	15.6	51.9
	(15.1)	(35.1)	(9.2)	(10.6)	(30.0)	(100.0)
Over 75	5.5	14.7	2.9	5.3	9.8	38.2
	(14.3)	(38.5)	(7.7)	(13.8)	(25.7)	(100.0)
			(Income group)			
Lowest 20%	4.9	11.9	2.5	2.2	7.2	28.8
	(16.9)	(41.4)	(8.8)	(7.7)	(25.1)	(100.0)
Second 20%	6.6	16.1	4.0	3.3	11.2	41.2
	(16.1)	(39.0)	(9.6)	(8.1)	(27.2)	(100.0)
Third 20%	8.5	20.7	5.6	3.8	17.4	55.9
	(15.1)	(37.0)	(10.0)	(6.8)	(31.1)	(100.0)
Fourth 20%	10.9	26.8	7.5	4.5	26.0	75.6
	(14.4)	(35.4)	(9.9)	(5.9)	(34.4)	(100.0)
Highest 20%	16.3	43.9	11.6	6.3	49.9	128.0
	(12.7)	(34.3)	(9.1)	(4.9)	(39.0)	(100.0)

Figures

Figure 1. Overview of the regional econometric input-output model (REIM)

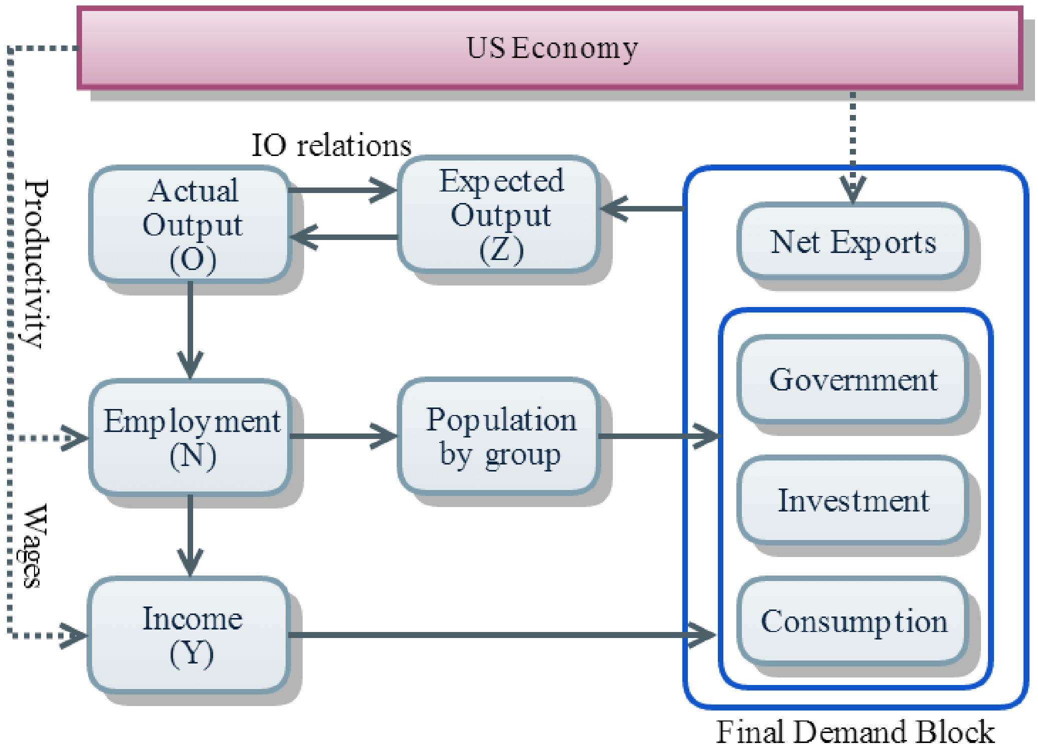


Figure 2. The 2011 total expenditure by age and income groups in the Chicago region

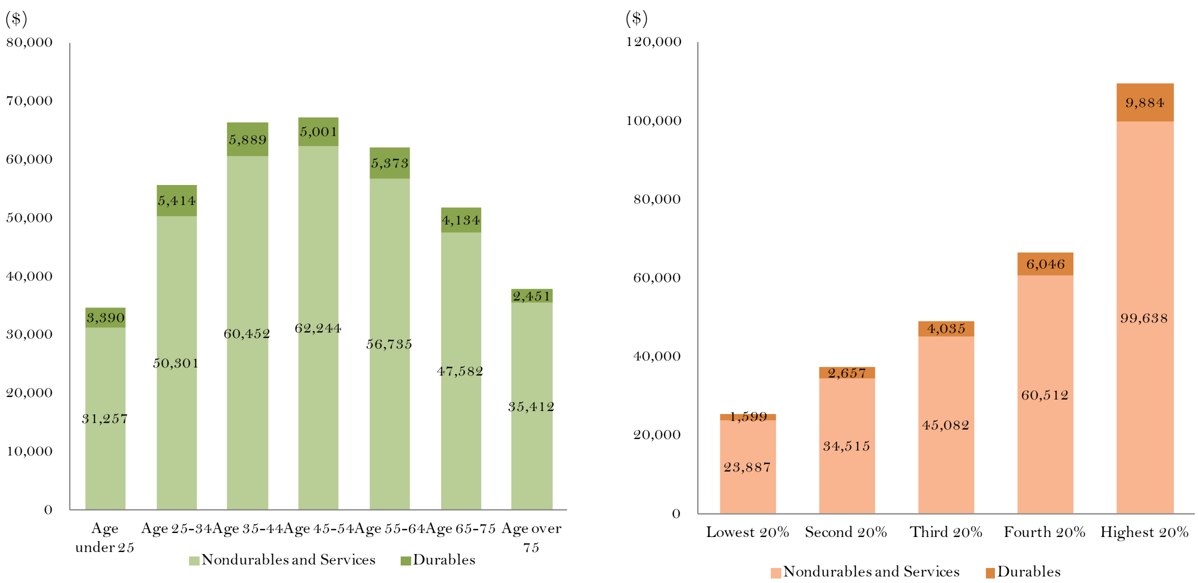


Figure 3. Average budget shares by age group (1984-2011)

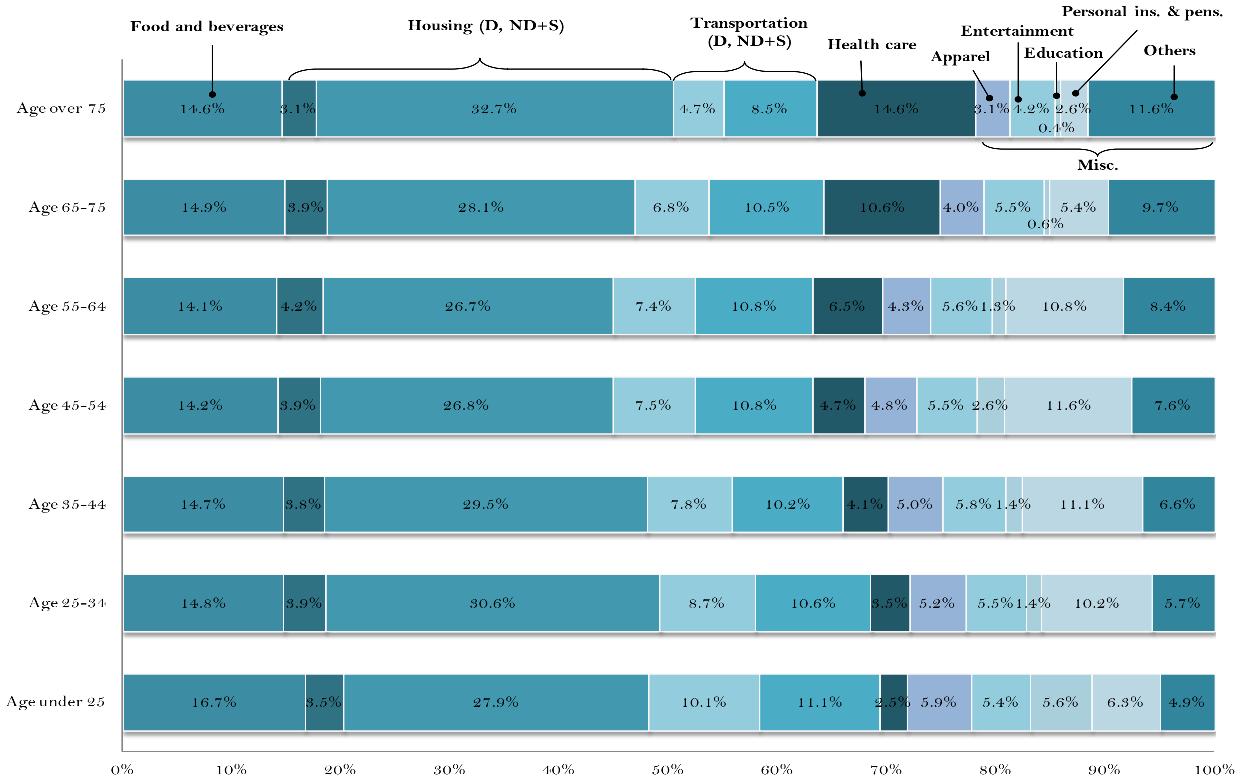


Figure 4. Average budget shares by Income quintile group (1984-2011)

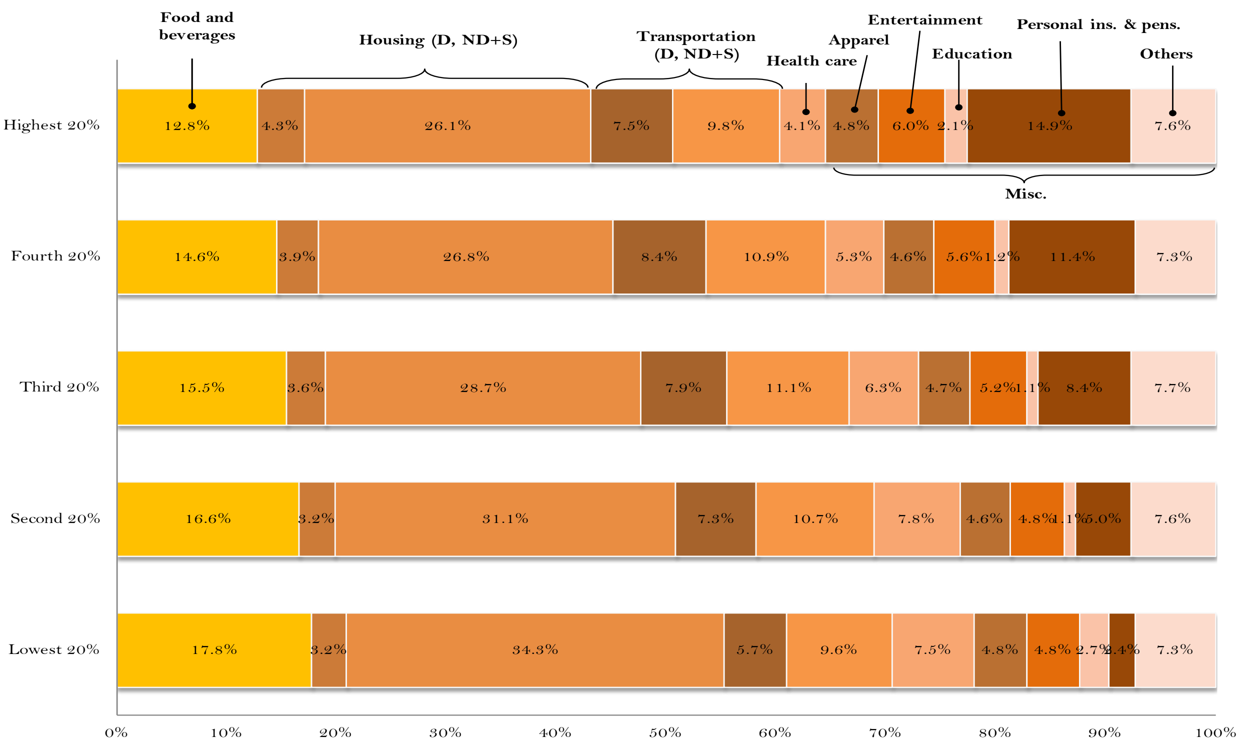


Figure 5. Average number in consumer unit by age group in the CES for the US

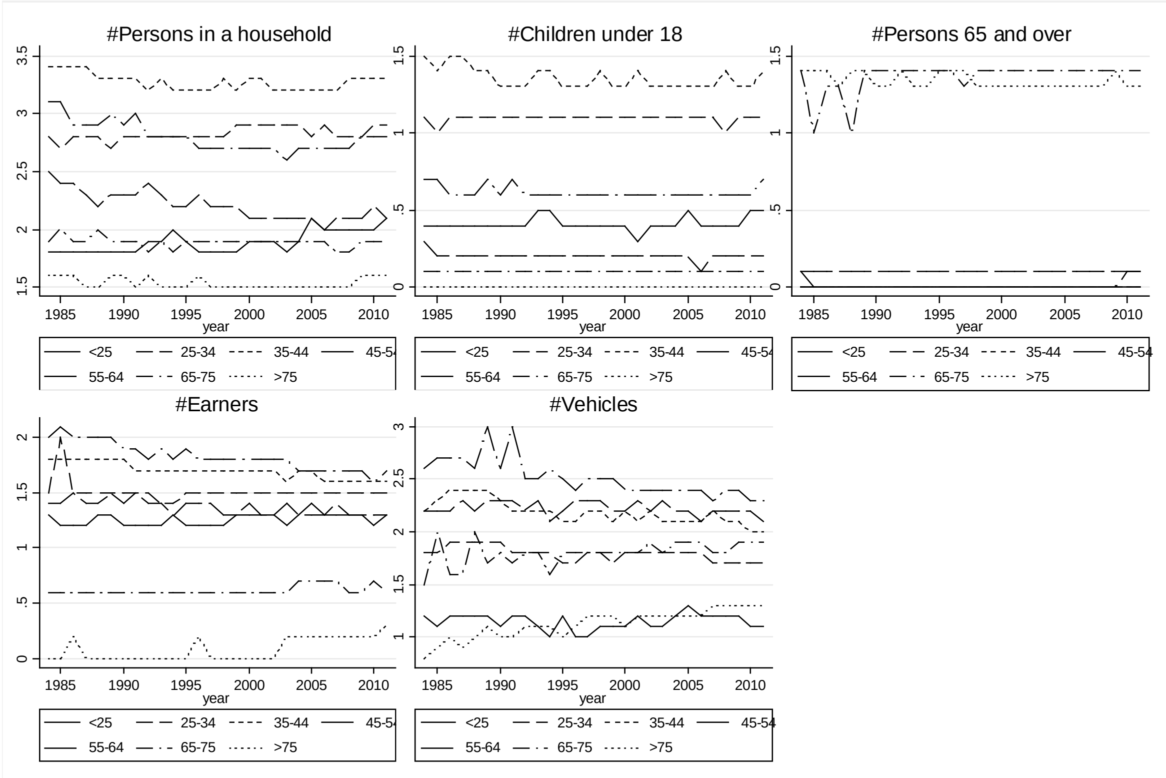


Figure 6. Average number in consumer unit by income group in the CES for the US

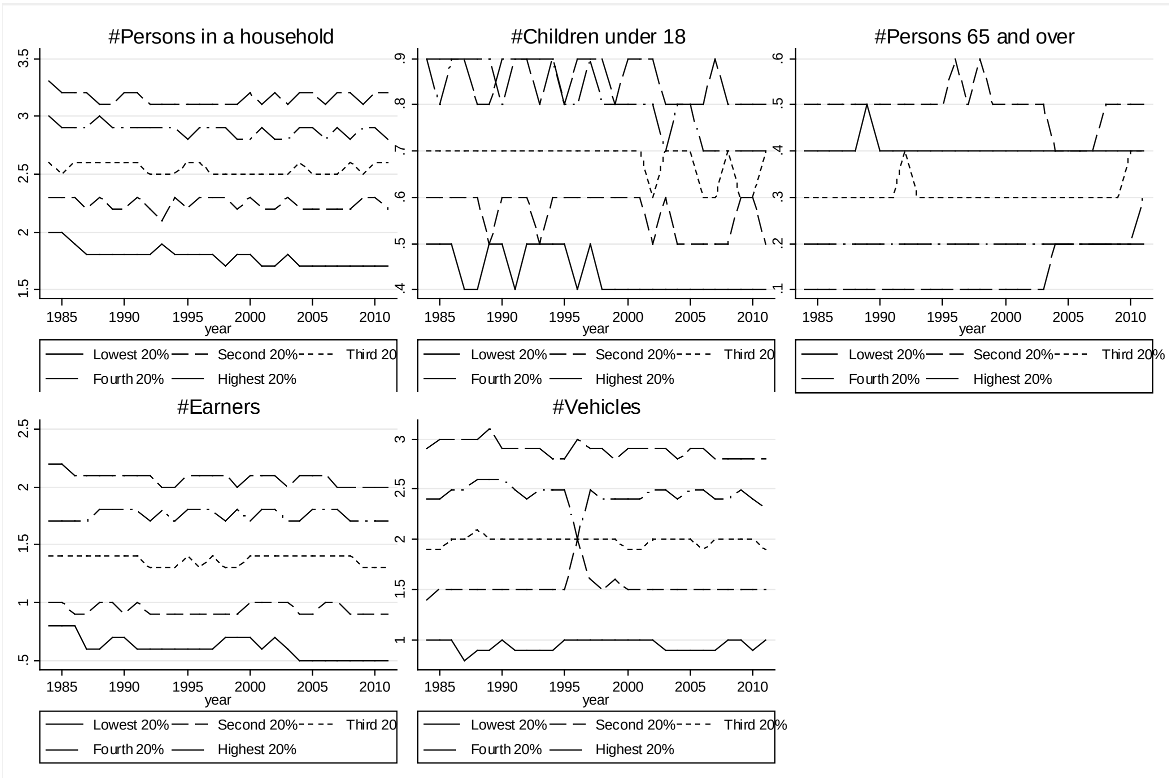
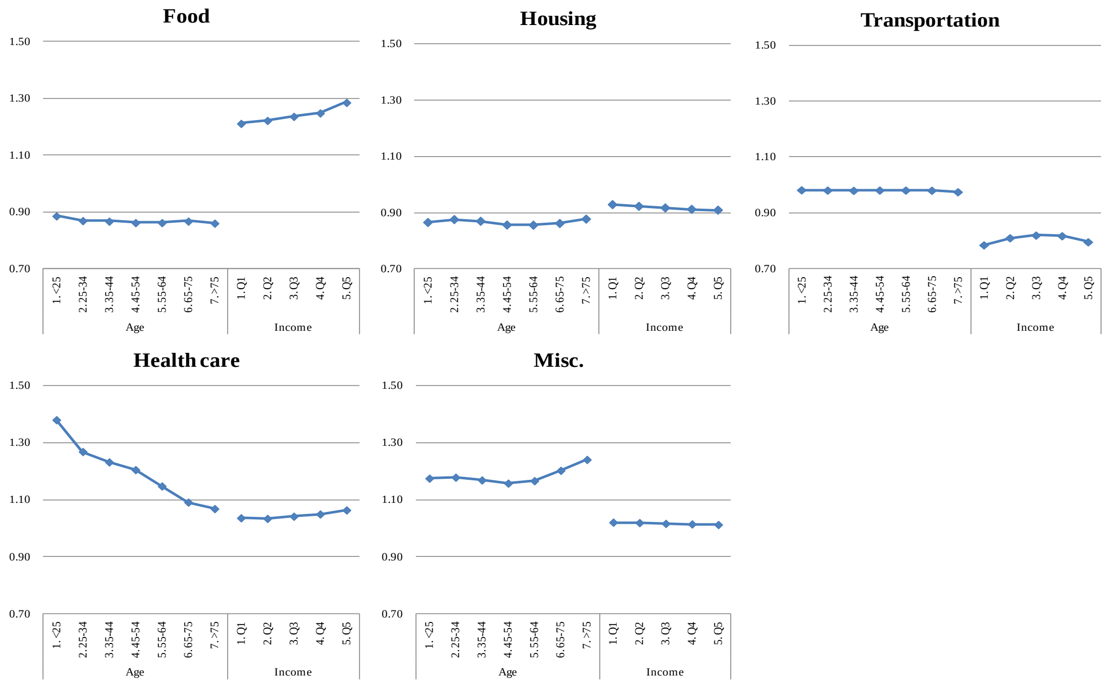


Figure 7. Estimated elasticities

(Total expenditure elasticities)



(Uncompensated own-price elasticities)

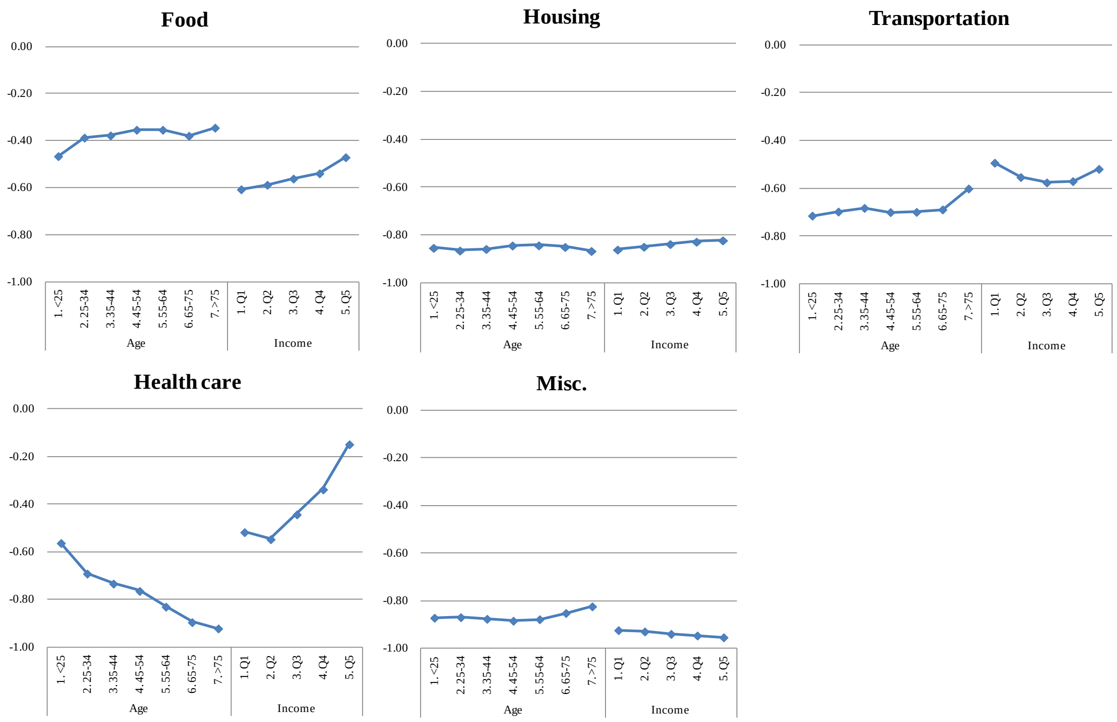
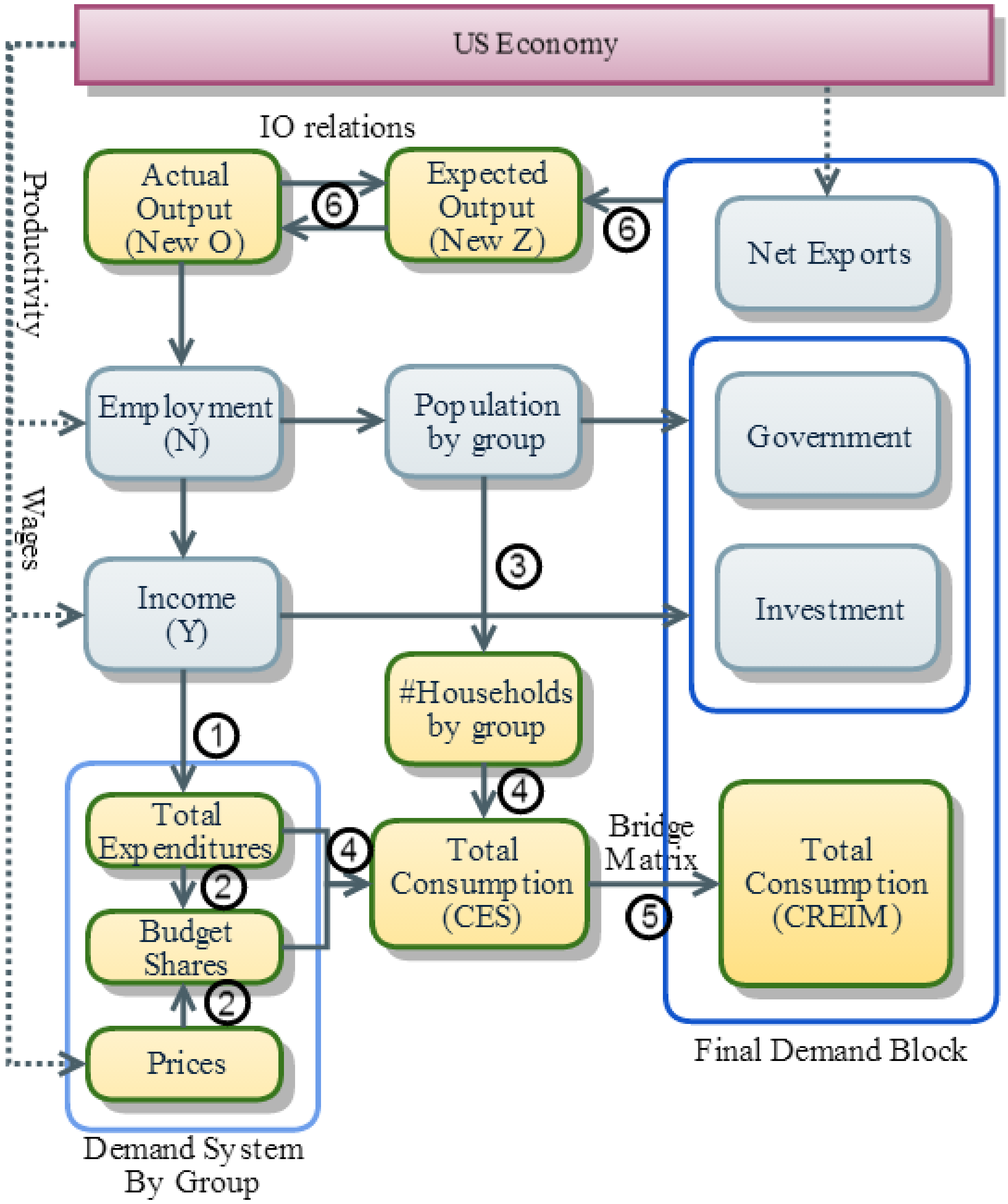


Figure 8. A schematic representation of the extended REIM



* Details on the circled numbers are described in text.

Figure 9. Outlook for expenditure shares by expenditure type by cohort (in current dollars)

(By Age) (By Income)

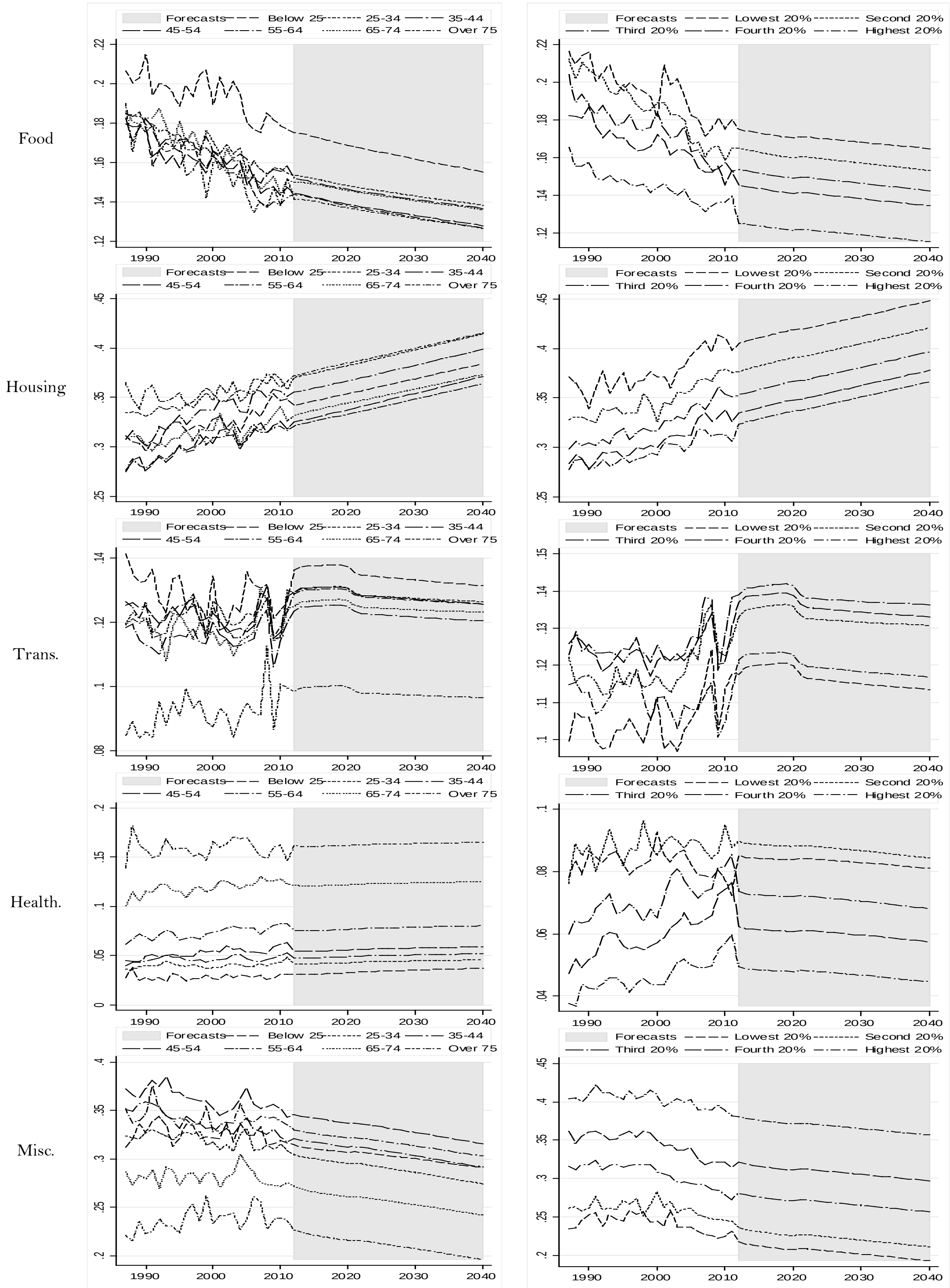
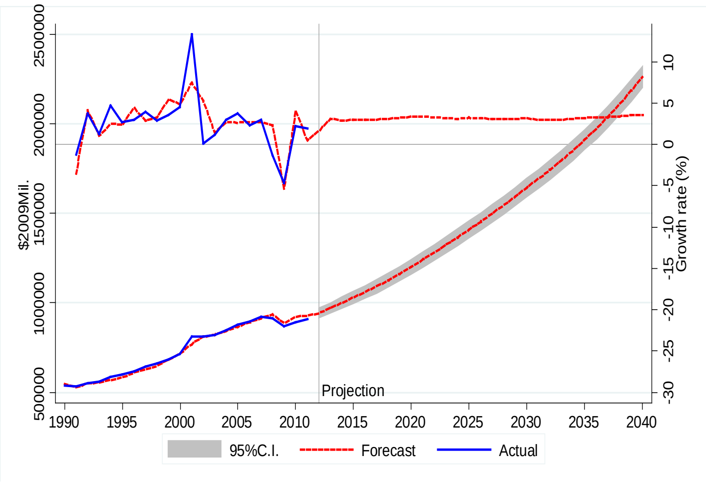
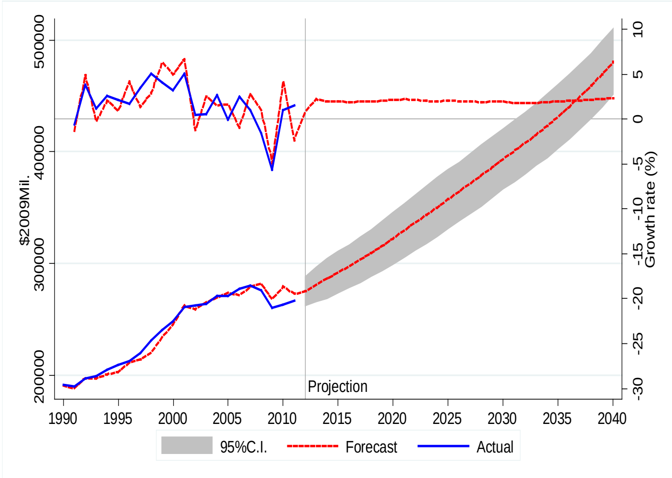


Figure 10. Baseline solutions for select endogenous variables

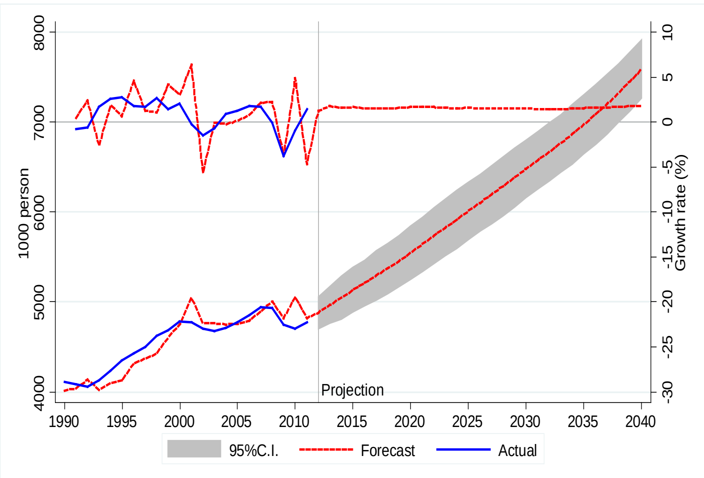
(Output)



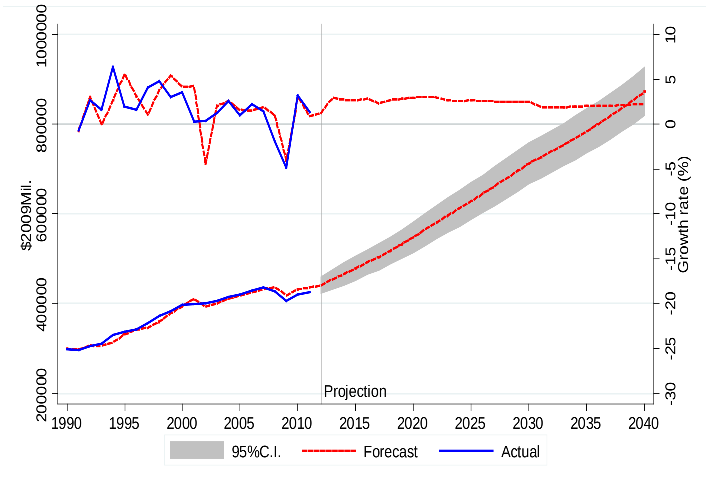
(Income)



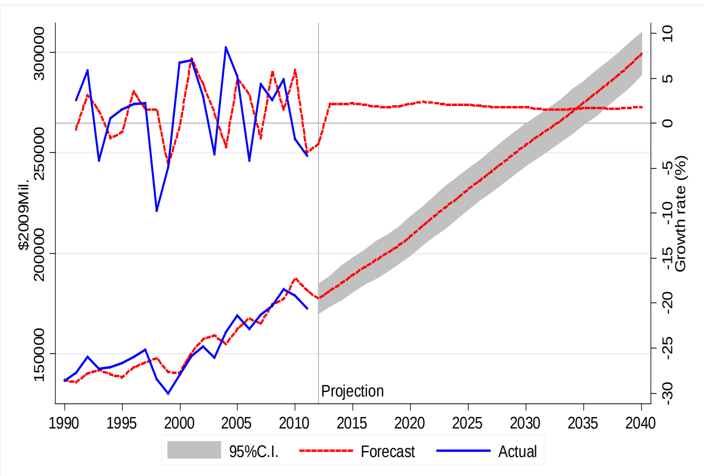
(Employment)



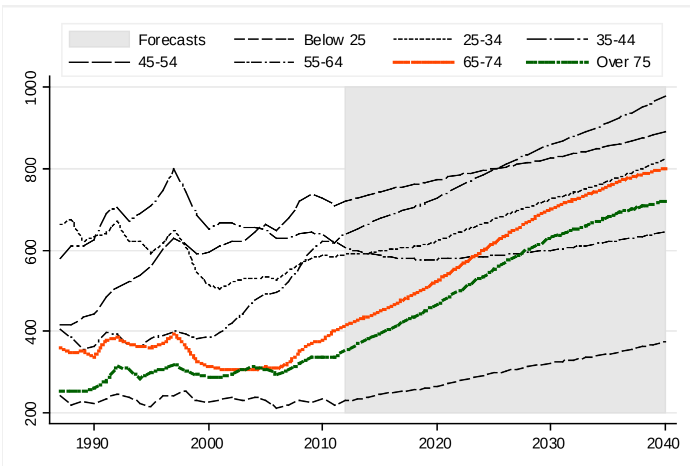
(Gross regional product)



(Consumption: Nondurables and services)

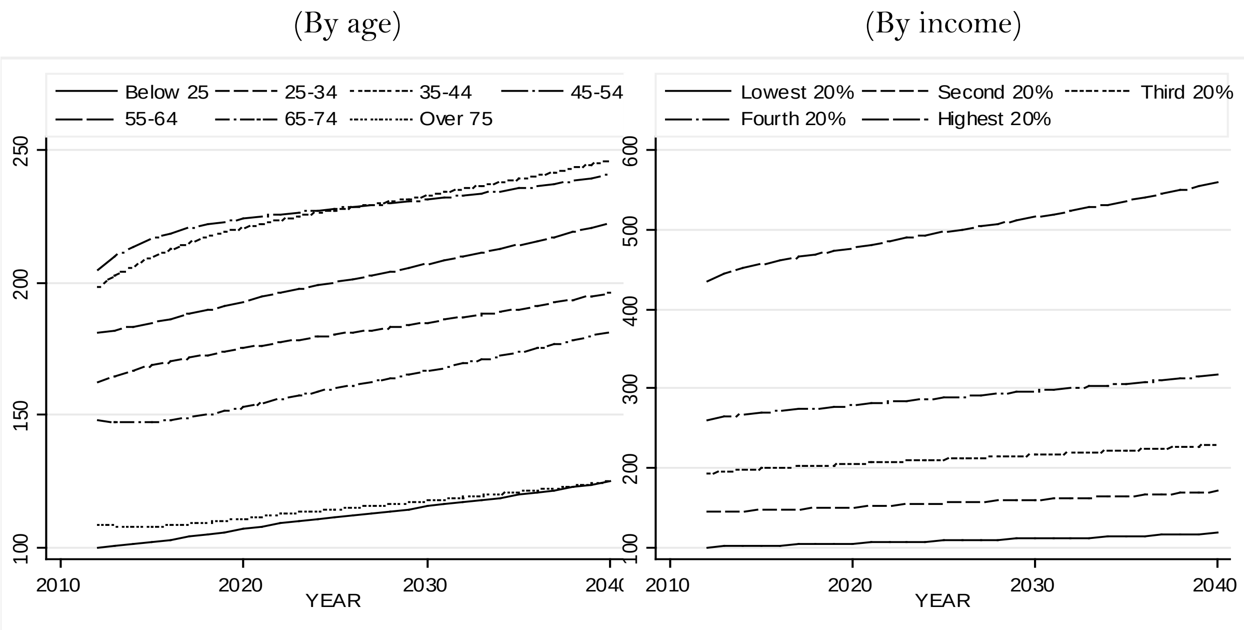


(Number of households by age)



* Unit: 1,000 household

Figure 11. Total impact of an 1,000 household increase by group on consumption over time



* Under 25 in 2012 = 100 ($38.2 bil.); Lowest 20% in 2012 =100 ($30.5 bil.)

Figure 12. Group-specific expenditure shares by type due to an 1,000 household increase