Two Perspectives on Preferences and Structural Transformation*

Berthold Herrendorf
Arizona State University

Richard Rogerson
Princeton University and NBER

Ákos Valentinyi
Cardiff University, Hungarian Academy of Sciences and CEPR

July 23, 2011

Abstract

We assess the empirical importance of income and relative price effects for structural transformation in the postwar US. We explain how there are two natural approaches to connect models of structural transformation to the data: sectors may be categories of final expenditure or value added; for example, the service sector may be the final expenditure on services or the value added from service industries. We estimate preferences for each approach and find that with final expenditure categories income effects are the dominant forces behind structural transformation whereas with value added categories relative price effects are dominant. We show how the input–output structure of the US economy can reconcile these different findings.

Keywords: income effects; preferences; relative price effects; structural transformation.

JEL classification: O11; O14.

*The authors thank Stuart Low and Ed Prescott for many helpful conversations about the topic of this paper. For comments and suggestions, the authors thank Paco Buera, Joe Kaboski, Todd Schoellman, and the conference participants at the SED meetings in Prague (2007) and Istanbul (2009), the conference of the Society for the Advancement of Economic Theory in Kos (2007), the RMM conference in Toronto (2008), the conference on Developments in Macroeconomics at Yonsei University (2009), as well as seminar participants at ASU, Bonn, the Deutsche Bundesbank, Humboldt University, the San Francisco FED, Southampton, Toronto, UCLA, University of Pennsylvania, Western Ontario, and York (Toronto). For financial support, Herrendorf thanks the Spanish Ministry of Education (Grant SEJ2006-05710/ECON) and Rogerson thanks both the NSF and the Korea Science Foundation (WCU-R33-10005). Hubert Janicki, Loris Rubini, and Paul Schreck provided able research assistance.

1 Introduction

The reallocation of resources across the broad economic sectors agriculture, manufacturing, and services is a prominent feature of economic development. Kuznets (1966) referred to this reallocation as the process of structural transformation and included it as one of the main stylized facts of development. Recent work shows that extending the standard one–sector growth model to incorporate structural transformation can be important for addressing a variety of substantive issues.[1] However, there is no consensus on what constitutes a reasonable specification of preferences in this class of multi-sector versions of the growth model. In particular, while households may reallocate expenditure shares across broad sectors because of either income effects or relative price effects, the relative empirical importance of these two forces for structural transformation has not been established.[2] This paper seeks to answer the question: What is an empirically reasonable specification of preferences in these models?

The first contribution of our paper is to point out that there is a fundamental conceptual ambiguity in answering this seemingly simple question. A simple static example will help to illustrate the issue. Assume a stand–in household that is endowed with a unit of time and has preferences given by $u(c_g, c_s)$ where c_g and c_s are consumption of goods and services, respectively. There are two sectoral production functions, $c_g = f^g(h_g)$ and $c_s = f^s(h_s)$ where h denotes some type of labor input. Even given the names of the two sectors, there are still two very different interpretations of this model. If one interprets the sectoral production functions as value added production functions, consistency dictates that the arguments of the utility functions are necessarily the value added components of final consumption. In this case, a cotton shirt represents consumption of both goods and services: raw cotton and processing from the goods sector and retail services from services sector. We will call this the consumption value added approach. Alternatively, one could interpret the commodities in the utility function as

[1]See, for example, Laitner (2000) and Gollin et al. (2002) for an application to early development, Messina (2006), Rogerson (2008) and Ngai and Pissarides (2008) for the evolutions of hours worked in Europe and the U.S., Duarte and Restuccia (2010) for productivity evolutions in the OECD, Caselli and Coleman (2001) and Herrendorf et al. (2011a) for regional convergence, and Bah (2008) for identifying problem sectors in poor countries. Other contribution to the literature on structural transformation include Echevarria (1997), Kongsamut et al. (2001), Ngai and Pissarides (2007), Acemoglu and Guerrieri (2008), and Foellmi and Zweimüller (2008). Herrendorf et al. (2011b) provide a review of this literature.

[2]For example, Kongsamut et al. (2001), assume that only income effects matter, whereas Baumol (1967) and Ngai and Pissarides (2007) assume that only relative price effects matter.

the final consumption purchases of the household, i.e., a cotton shirt represents consumption of goods, while health care, for example, represents consumption of services. In this case consistency requires that the sectoral production functions be final consumption production functions rather than value added production functions.[3] We call this the final consumption expenditure approach. For a given model, the empirically reasonable choices for the parameters of utility and production functions will in principle depend on the choice of interpretation.[4] Failure to recognize this can lead to inconsistent specifications of utility and production functions in a given model, or to apparent inconsistencies in estimates across different empirical studies.

The second contribution of the paper is to provide empirically reasonable preference specifications to be used in models of structural transformation, for each of the interpretations just noted. Whereas the relevant data for the final expenditure approach is readily available, this is not true for the consumption value added approach. To be sure, data on total value added by sector are readily available, but these data are not sufficient because not all of total value added is consumed. One of the by–products of this paper is to lay out and implement a procedure for extracting the consumption component of total value added, and to produce an annual time series for U.S. consumption value added by sector between 1947 and 2007.

For each approach we find that a relatively simple utility function provides a good fit to the relevant data. However, the two specifications have fundamentally different economic implications, thereby emphasizing the empirical significance of the ambiguity noted above. For the final consumption expenditure approach, we find that a Stone–Geary utility function provides a good fit to these data. Moreover, we show that the nonhomotheticicities implicit in this specification are key in shaping the evolution of expenditure shares over the period 1947–2007, implying that income effects rather than relative price effects are the dominant force behind changes in expenditure shares. In terms of the recent theoretical literature on structural transformation, this result supports the specification assumed by Kongsamut et al. (2001), but not the one assumed

[3] Valentinyi and Herrendorf (2008) showed how to construct sectoral production functions that use only capital and labor to produce final expenditure by broad category.

[4] In the spirit of Lancaster (1966), one could adopt the view that households value a large set of characteristics that are bundled in various combinations in different goods. In general, any attempt to capture this complex ordering using a utility function with two arguments will lead to some distortions. In our view these two different specifications each have their individual strengths and weaknesses in terms of capturing relevant aspects of preferences. We therefore do not think that it is useful to identify one of them as the preferred approach.

by Baumol (1967) and Ngai and Pissarides (2007). For the consumption value added approach we find that nonhomotheticities are relatively unimportant and that a homothetic specification with no substitution between categories (i.e., a Leontief specification) provides a good fit to the data. In other words, contrary to our earlier results, it is now the preference specification adopted by Baumol (1967) and Ngai and Pissarides (2007), rather than the one adopted by Kongsamut et al. (2001), that is consistent with the data. With these preferences, the dominant force behind changes in expenditure shares is relative price effects, rather than income effects.

These results should not be interpreted as saying that researchers can freely rationalize very different choices of utility functions. Our two estimated utility functions are based on two different representations of the same underlying data, instead of two different data sets. The final consumption expenditure data are linked to the consumption value added data through complicated input–output relationships, which implicitly translate the income effects that dominate with final consumption expenditure into the relative price effects that dominate with consumption value added, and vice versa. That is, one could interpret the production functions f^g and f^s in our earlier example as value added production functions and then specify two input–output functions that map value added from the goods and services sectors into the final consumption expenditure of goods and services. The third contribution of our paper is to explore how the input–output structure influences the mapping between the two different representations. We show analytically for one special case that the non–homotheticities from the final expenditure specification are eliminated in the value added specification. While estimation of the input–output structure rejects this special case statistically, we show that reality is still quite close to it, therefore reconciling our different findings for the two specifications.

The key point to be taken away is that in the context of multi–sector general equilibrium models, one needs to be careful about how commodities are defined, and in particular about consistency of measurement on the household and production sides of the economy. Moreover, one needs to be careful about importing parameter estimates across models that may have different underlying definitions of commodities, even though they may label them in the same way. For example, our results show that it is not appropriate to use the utility function that was estimated from final consumption expenditure together with value added production functions at the sector level. If one wants to use a utility function that was estimated from final consumption expenditure, then one either needs to write down a production structure that captures the complexities of the input–output relationships at the sector level, or one needs to find a representation of production that isolates the contribution of capital and labor to the production of final expenditure categories. While this can be done, it is much more difficult than working directly with sectoral value added production functions.[5]

An outline of the paper follows. In the next section we describe the model and the method that we use to calibrate preference parameters. In Section 3 we describe the final consumption expenditure method and we report the estimation results for this method. In Section 4, we turn to consumption value added. We explain in some detail how to construct the relevant time series of variables from existing data and we report the estimation results. Section 5 links the results of both methods and provides intuition for the differences. Moreover, it discuss the relative merits of the two methods and some additional measurement issues. Section 6 concludes. An appendix contains the details about our data work.

2 Model

As noted in the introduction, our objective is to determine what form of preferences for a stand–in household defined over broad categories are consistent with U.S. data for expenditure shares over the period since 1947. This section develops the model that we use to answer this question.

We consider an infinitely lived household with preferences represented by a utility function of the form:

$$\sum_{t=0}^{\infty} \beta^t U(c_{at}, c_{mt}, c_{st}, 1 - h_t),$$

where the indices a, m, and s refer to the three broad sectors of agriculture, manufacturing, and services.[6] Hours of work for the household in period t are denoted by h_t and, with the total time

[5] Valentinyi and Herrendorf (2008) showed how to construct sectoral production functions that use only capital and labor to produce final expenditure by broad category.

[6] The exact definition of these sectors for each of the two specifications that we consider will be provided later. We note here that we have followed the convention of using the label "manufacturing" to describe a sector which consists of manufacturing and some other sectors (e.g., mining and construction). While the label "industry" is sometimes used to describe this sector, we will later use the term "industry" to describe a generic production activity and the index i to denote a generic sector. In view of this, "manufacturing" seemed a better choice.

endowment normalized to 1, the term $1 - h_t$ represents leisure in period t. We will assume that the function U takes one of two forms:

$$U(c_{at}, c_{mt}, c_{st}, 1 - h_t) = \frac{u(c_{at}, c_{mt}, c_{st})^{1-\rho} - 1}{1 - \rho} v(1 - h_t)$$

or

$$U(c_{at}, c_{mt}, c_{st}, 1 - h_t) = \frac{u(c_{at}, c_{mt}, c_{st})^{1-\rho} - 1}{1 - \rho} + v(1 - h_t).$$

where $\rho > 0$ is the intertemporal elasticity of substitution of consumption. The key feature of these two forms is that time devoted to work has no effect on relative marginal utilities of consumption within a given period, so it will not influence the optimal allocation of expenditures across consumption categories for given prices and total expenditure.

We further assume that the period utility function $u(c_{at}, c_{mt}, c_{st})$ is of the form:

$$u(c_{at}, c_{mt}, c_{st}) = \left(\sum_{i=a,m,s} \omega_i^{\frac{1}{\sigma}} (c_{it} + \bar{c}_i)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{1}$$

where the ω_i are non–negative weights that add up to one and $\bar{c}_i$ are constants. We restrict $\bar{c}_m$ to be zero but allow $\bar{c}_a$ and $\bar{c}_s$ to take any value.[7] If the $\bar{c}_i$'s are all zero, then preferences are homothetic and $\sigma > 0$ is the within period elasticity of substitution between consumption goods.

This utility specifications nests both Kongsamut et al. (2001) and Ngai and Pissarides (2007). The preferences used by Kongsamut et al are the special case in which $\sigma = 1$, $\bar{c}_a < 0$ and $\bar{c}_s > 0$. The implied utility function was first introduced by Stone (1954) and Geary (1950-1951):[8]

$$u(c_{at}, c_{mt}, c_{st}) = \omega_a \log(c_{at} + \bar{c}_a) + \omega_m \log(c_{mt}) + \omega_s \log(c_{st} + \bar{c}_s). \tag{2}$$

The preferences used by Ngai and Pissarides are the special case in which $\sigma < 1$ and $\bar{c}_a = \bar{c}_s =$

[7] We follow Kongsamut et al. (2001) in restricting $\bar{c}_m$ to equal zero. We have experimented with an unrestricted specification where $\bar{c}_m$ could take any value but found that the goodness of fit hardly changed.

[8] The implied demand model is often called the Linear Expenditure System. Deaton and Muellbauer (1980) is another classic contribution to the literature on expenditure systems.

0. This is a homothetic CES specification with less substitutability than log:

$$u(c_{at}, c_{mt}, c_{st}) = \left(\sum_{i=a,m,s} \omega_i^{\frac{1}{\sigma}} c_{it}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{3}$$

Note that we assumed that σ is the same among all three consumption categories, which may seem somewhat restrictive. Instead, we could have considered more general utility functions with two levels of CES aggregators, each with (possibly) different elasticities of substitution.[9] We did not pursue this possibility further because we found that our parsimonious class of utility functions includes specifications that provide a very good fit to the expenditure shares. Note also that if all individuals have preferences of the above form and have total consumption expenditure that exceed a minimum level, then aggregate expenditures are consistent with those for a stand–in household with preferences of the same form.[10] This property extends to settings in which individuals make consumption/savings decisions if there are complete markets.

Consider the stand–in household in a setting in which it maximizes lifetime utility given a market structure that features markets for each of the three consumptions, a labor market, and a market for borrowing and lending at each date t.[11] Our strategy is to focus solely on the implications for optimal consumption behavior within each period. The advantage of this "partial" approach is that we do not have to take a stand on the exact nature of intertemporal opportunities available to the household (i.e., the appropriate interest rates for borrowing and lending), or to specify how expectations of the future are formed. With these assumptions, if C_t is observed total expenditure on consumption in period t and p_{it} are observed prices, then it follows that the consumption choices in period t must solve the following static optimization problem:

$$\max_{c_{at}, c_{mt}, c_{st}} u(c_{at}, c_{mt}, c_{st}) \quad \text{s.t.} \quad \sum_{i=a,m,s} p_{it} c_{it} = C_t.$$

Assuming interior solutions, the first–order conditions for the above maximization problem

[9] See Sato (1975) for a characterization of general CES utility functions.

[10] The precise condition is (13) in Appendix A where we derive this result formally.

[11] More generally, we could assume some uncertainty in the economy and allow for a richer set of assets that can be traded; see for example Atkeson and Ogaki (1996). What matters for our method is that there are spot markets at each date t for each of the three consumption categories.

are easily derived.[12] Some simple algebra yields the following expression for the expenditure shares:

$$s_{it} \equiv \frac{p_{it}c_{it}}{C_t} = \frac{1 + \sum_{j=a,m,s} \frac{p_{jt}\bar{c}_j}{C_t}}{\sum_{j=a,m,s} \frac{\omega_j}{\omega_i}\left(\frac{p_{jt}}{p_{it}}\right)^{1-\sigma}} - \frac{p_{it}\bar{c}_i}{C_t}. \tag{4}$$

In the empirical worked reported below, we will estimate the parameters of the utility function employing the nonlinear seemingly unrelated regression framework, which here is equivalent to maximum likelihood.[13]

3 Final Consumption Expenditure

The final consumption expenditure method originated in the literature on expenditure systems. It associates the arguments of the utility function with final expenditure of households over different categories of goods and services. Specifically, this method classifies the expenditures on individual *commodities* into the three broad sectors of agriculture, manufacturing, and services. For example, purchases of food from supermarkets will be included in c_{at}, purchases of clothing will be included in c_{mt}, and purchases of air–travel services will be included in c_{st}.

3.1 Implementing the Final Consumption Expenditure Specification

The required data in this case are total consumption expenditure and the expenditure shares and prices for final consumption expenditure on different commodities. These data are readily available from the BEA.[14]

While expenditure shares do not depend on how one splits total expenditures into their price and quantity components, the series for prices do. That is, given total expenditure, different procedures for inferring the consumption quantities will imply different relative prices. Consistent with BEA measurement, we measure final consumption quantities using chain–weighted meth-

[12]In general, of course, the nonhomotheticity terms in our class of utility functions can lead to corner solutions. However, this is not relevant for *aggregate* consumption in a *rich* country such as the postwar U.S.A. Looking ahead, indeed we will find that the stand–in household chooses quantities that are far away from corners.

[13]For further discussion on the econometric issues, see the review of Barnett and Serletis (2008).

[14]The exact data sources can be found in Appendix B.1.

ods. For the period 1947–2007 and for the available commodities, we obtain annual data on final consumption expenditure, chain–weighted final consumption quantities, and prices from the BEA. Since quantities calculated with the chain–weighted method are not additive, we use the so called cyclical expansion procedure to aggregate quantities that are not available from the BEA.[15] We assign each commodity to one of the three broad sectors agriculture, manufacturing, and services. A detailed description of this assignment can be found in Appendix B.2. Note that for estimating utility function parameters we do not need to know whether the commodities purchased by the household are produced in the U.S. economy or imported. All that matters for our exercise is information on total consumption expenditure, expenditure shares and prices.

Figures 1–3 show the resulting evolution of the expenditure shares, prices and quantities, respectively. Looking at Figure 1, we see that the data are consistent with the standard (asymptotic) pattern of structural transformation: The expenditure share for services is increasing, while those for agriculture and manufacturing are decreasing. Turning next to Figure 2, which shows the evolution of prices (with prices in 1947 normalized to 1), we see that while all three prices have increased, the price of services has increased relative to both manufacturing and agriculture and the price of agriculture has increased relative to manufacturing. Figure 3 shows real quantities relative to their 1947 values. Here we see that while the quantities of all three categories have increased, the quantity of manufacturing has grown the most, while the quantity of agriculture has grown the least.

Figures 1–3 already suggest some of the qualitative features of the utility specification that our estimation will select. First, note that the price of services has increased relative to that of agriculture, while at the same time the quantity of services has also increased substantially relative to that of agriculture. This is qualitatively inconsistent with a homothetic utility specification, which would have relative prices and relative quantities move in opposite directions. In the context of our class of utility functions, reconciling these observations amounts to having $\bar{c}_a < 0$ and/or $\bar{c}_s > 0$. Second, as the price of agriculture relative to manufacturing has increased, the quantity of agriculture relative to manufacturing has decreased. This is consistent with there being substitutability between agriculture and manufacturing. While to some extent this could

[15] See Landefeld and Parker (1997) for more details.

Final Consumption Expenditure Per Capita

Figure 1: Expenditure shares

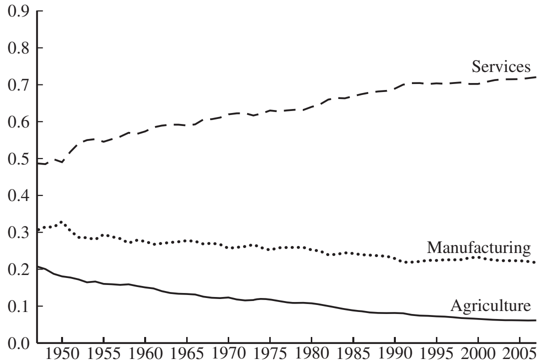


Figure 2: Price Indices (1947=1)

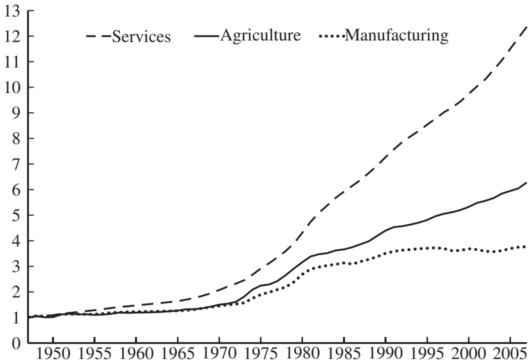
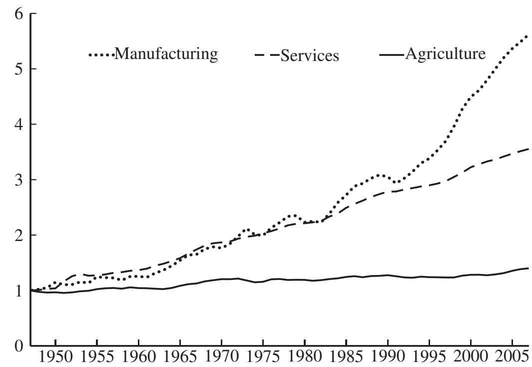


Figure 3: Quantity Indices (2000 chained dollars, 1947=1)



also be accounted for by having $\bar{c}_a < 0$, in the context of our preference specification, it turns out that σ will come out close to one.

3.2 Results with Final Consumption Expenditure

We now estimate the parameters of the utility function by estimating the three demand functions expressed as expenditure shares in equation (4). We use an iterated feasible generalized nonlinear least square estimator, which is equivalent to maximum likelihood in our setting. Since the expenditure shares sum to one, the error covariance matrix is singular. Therefore we drop the demand for agricultural goods when we do the estimation. Note that the estimation results are not affected by which equation we drop. To deal with the issue that four out of our six parameters are constrained (i.e., $\sigma \geq 0$, $\omega_a + \omega_m + \omega_s = 1$ with $\omega_i \geq 0$) we transform the constrained parameters into unconstrained parameters in the following way:

$$\sigma = e^{b_0}, \quad \omega_a = \frac{1}{1 + e^{b_1} + e^{b_2}}, \quad \omega_m = \frac{e^{b_1}}{1 + e^{b_1} + e^{b_2}}, \quad \omega_m = \frac{e^{b_2}}{1 + e^{b_1} + e^{b_2}},$$

where $b_0, b_1, b_2 \in (-\infty, +\infty)$. We estimate the model in terms of the unconstrained parameters b_0, b_1, b_2 and $\bar{c}_a, \bar{c}_s$ and then calculate the point estimates and standard errors of the constrained parameters $\sigma, \omega_a, \omega_m, \omega_s$.

Table 1 shows the results for three different specifications. For now we focus on the first two columns; the estimates from the third column are discussed in the next subsection. Column (1) shows the results when we do not impose any restrictions on the parameters other than $\bar{c}_m = 0$. The implied value of σ is 0.94 and the signs of the two unrestricted nonhomothetic terms have the pattern that Kongsamut et al. (2001) suggested, that is, $\bar{c}_a < 0$ and $\bar{c}_s > 0$. The nonhomotheticities imply that as total expenditure grows holding relative prices fixed, the consumption share of agriculture will go down while that of services will go up.[16] Figure 4 shows that the fit of the estimated model from Column (1) to the data on final consumption expenditure shares is very good.

While the specification from Column (1) is similar to the Stone–Geary specification im-

[16] This result is broadly consistent with what panel studies of household consumption find; see Houthakker and Taylor (1970) for a classic contribution.

Table 1: Results with Final Consumption Expenditure

	(1)	(2)	(3)
σ	0.94^{**}	1	0.29^{**}
	(0.06)	–	(0.04)
$\bar{c}_a$	-1175.73^{**}	-1163.31^{**}	
	(28.87)	(24.03)	
$\bar{c}_s$	12867.89^{**}	16056.46^{**}	
	(3508.55)	(997.10)	
ω_a	0.02^{**}	0.02^{**}	0.11^{**}
	(0.0002)	(0.001)	(0.004)
ω_m	0.16^{**}	0.15^{**}	0.20^{**}
	(0.01)	(0.004)	(0.004)
ω_s	0.82^{**}	0.83^{**}	0.69^{**}
	(0.02)	(0.004)	(0.003)
AIC	-895.44	-897.04	-599.92
$RMSE_a$	0.004	0.004	0.033
$RMSE_m$	0.009	0.009	0.018
$RMSE_s$	0.010	0.011	0.023

Robust standard errors in parentheses * $p < 0.05$, ** $p < 0.01$
AIC is the Akaike information criterion, $RMSE_i$ is the root mean squared error for equation i.

Fit of Expenditure Shares with Final Consumption Expenditure

Figure 4: Fit of Column (1)

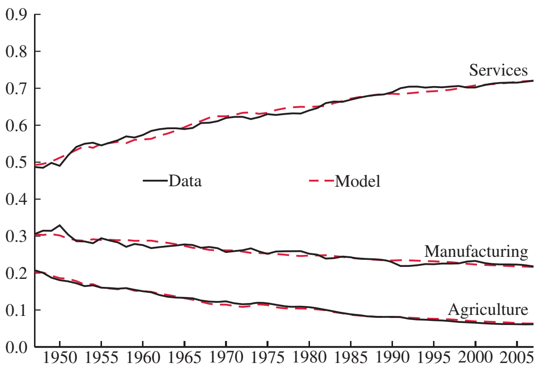
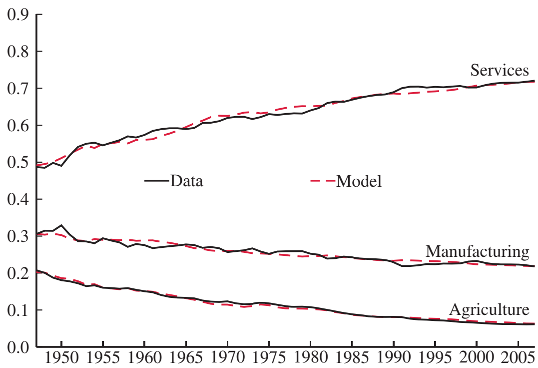


Figure 5: Fit of Column (2)



posed by Kongsamut et al. (2001), it is not identical, since they assumed $\sigma = 1$. It is therefore interesting to assess the extent to which the specification of Kongsamut et al. (2001) fits the data. To examine this, the second column redoes the estimation, this time imposing that $\sigma = 1$. The nonhomothetic terms retain the same sign configuration, although the magnitude of $\bar{c}_s$ changes significantly. This is intuitive because with a higher σ households react to the large increase in the relative price of services by substituting away from services. The higher $\bar{c}_s$ term compensates for this. Figure 5 shows that the specification of Column (2) fits virtually as well as the specification of Column (1). This is consistent with the fact that the Akaike information criterion (AIC) in Table 1 hardly changes.[17]

Several papers have previously estimated the linear expenditure systems that are implied by the Stone–Geary utility specification. Our results are most closely related to those of a literature that used time series data for final consumption expenditure per capita. A prominent example is Pollak and Wales (1969), who used US data 1948-1965 on the four broad categories food, clothing, shelter, and miscellaneous. Pollak and Wales also found that the linear expenditure system fits the data very well and that nonhomotheticity terms are important.[18]

We conclude that when using data on final consumption expenditure, the data broadly support the Stone–Geary specification of Kongsamut et al. (2001). Having said that, note that they also imposed the condition

$$p_{at}\bar{c}_a + p_{st}\bar{c}_s = 0,$$

which is required for the existence of a generalized balanced growth path in their model.[19] This condition is rather trivially not consistent with the data we use. The simplest way to see this is to look at Figure 2, which clearly shows that p_{st}/p_{at} has been steadily increasing since 1947, whereas, of course, $\bar{c}_a$ and $\bar{c}_s$ are constants.

[17] We do not report the standard R^2 statistic here because it is not well defined for non–linear regressions. Instead, we report the Akaike information criterion and the root mean squared errors here. Note that in itself the value of the Akaike information criterion does not convey information. What conveys information is whether that value increases (improved fit) or decreases (reduced fit) if one goes from one specification to another.

[18] For a subsequent literature review, see Blundell (1988).

[19] Given the nonhomotheticity terms, their model does not have a balanced growth path in the usual sense of the word. They therefore consider a generalized balanced growth path, which they define as a growth path along which the real interest rate is constant.

Table 2: Nonhomotheticity Terms Relative to Final Consumption Expenditure from the Data

	1947	2007
$-p_a\bar{c}_a/C$	0.16	0.03
$p_s\bar{c}_s/C$	0.98	0.41
$-\bar{c}_a/c_a$	0.79	0.57
$\bar{c}_s/c_s$	2.02	0.57

3.3 Income versus Price Effects with Final Consumption Expenditure

It is of interest to look more closely at the importance of income versus relative price effects in accounting for the observed changes in the shares of final consumption expenditures. To put the size of the estimated nonhomotheticity terms of Column (1) into perspective, Table 2 reports the $\bar{c}_i$'s relative to final consumption expenditure from the data. Most importantly, rows three and four show that in 1947 both nonhomotheticity terms were sizeable compared to the actual consumption quantities of agriculture and services. This suggests that income effects play a very important role in shaping the shares of final consumption expenditure.

To explore further the importance of income versus relative price effects, Figure 6 shows the fit of the expenditure shares implied by the parameters of Column (1) when total expenditure changes as dictated by the data but relative prices are held constant at their 1947 values. We can see that although the fit deteriorates somewhat, the model still captures the vast majority of the changes in the expenditure shares. The main discrepancies between the data and the model are that the share of services now increases somewhat more than in the data and the share of agriculture now decreases somewhat more than in the data. These discrepancies are intuitive since the price of services increases relative to agriculture during our sample period, and would therefore work to partially offset the changes associated with the income effects.

A second way of judging the importance of income versus relative price effects is to ask how well a homothetic specification can fit the data, since such a specification necessarily implies that total expenditure has no effect on expenditure shares. Column (3) of Table 1 presents the results of a specification in which the nonhomothetic terms are restricted to equal zero. We

Income Effects with Final Consumption Expenditure Shares

Figure 6: Fit of Column (1) with Relative Prices Fixed at 1947 Values

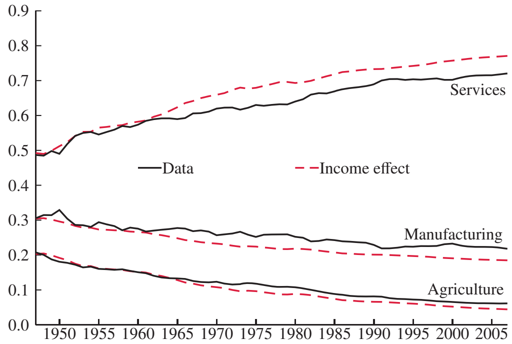
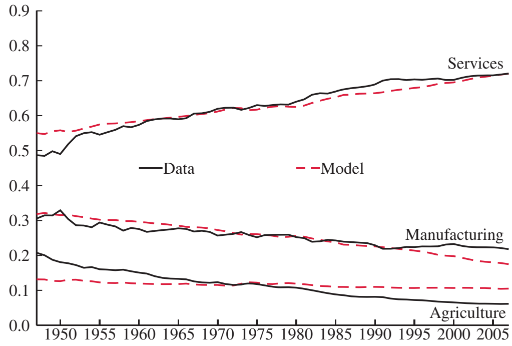


Figure 7: Fit of Homothetic Specification in Column (3)



can see that the estimated elasticity term σ drops from 0.94 to 0.27. Moreover, the Akaike information criterion significantly increases, suggesting that the fit gets considerably worse. Figure 7 confirms this by showing that the fit becomes fairly bad for agriculture relative to the previous two specifications. Absent income effects and given a rising relative price of services, the way in which the homothetic specification can account for the large rise in the expenditure share for services is by having very little substitutability. But, as noted earlier, absent income effects, fitting the expenditure share for agriculture requires some substitutability. Hence, the model without income effects cannot do a good job of matching all expenditure shares.

We conclude that the income effects associated with the nonhomotheticities are the dominant source of the observed structural transformation in the shares of final consumption expenditure.

4 Consumption Value Added

As noted in the introduction, many multi–sector general equilibrium models represent the sectoral production functions in value added form, in which case the arguments of the utility function necessarily represent the value added components of final expenditure. Individual *industries* are then classified into the three broad sectors agriculture, manufacturing, and services, and a sector is a collection of industries, with sector value added being the sum of the value added of the industries belonging to it. Effectively, this way of proceeding breaks consumption spending into its value added components. For example, purchases from supermarkets will then be broken down into the components of c_{at} (food), c_{mt} (processing of the food) and c_{st} (distribution services). Similarly, purchases of clothing will be broken down into the components of c_{at} (raw materials, say cotton), c_{mt} (processing of cotton into clothing) and c_{st} (distribution services), and purchases of air–travel services will be broken down into the components of c_{mt} (fuel) and c_{st} (transportation services).

Whether one prefers to use final consumption expenditure or consumption value added will depend on data availability and the specific application. While we will discuss the relative merits of each method in more detail in subsection 5.2 below, there are two key points that we want to emphasize here already. First, there is no reason to believe that the parameters of the utility function are invariant to the definition of commodities. Looking ahead, the distinction between the two different specifications will turn out to be very significant, since we will find that they imply preferences with very different qualitative properties. Second, and related, it is important to emphasize that the two specifications are two different representations of the same underlying data. Put differently, the data on final consumption expenditure are linked to the data on consumption value added through complicated input–output relationships, and vice versa. Our results should therefore not be interpreted as implying that different data sets provide different parameter estimates. Rather, they tell us that different transformations of given data have different properties. We explore the mapping between these two specifications in more detail in a later section.

4.1 Implementing the Consumption Value Added Specification

We now describe how to construct the relevant data when one identifies the three consumption categories with their respective value added components. The exact data sources can be found in Appendix B.1. Similar to the case of final expenditure shares, there is annual data available from the BEA on value added by industry, as well as real value added and prices. As we mentioned above, the consumption value added method assigns industries, instead of commodities, to the three broad sectors. Appendix B.2 describes the details of this assignment.

Consumption Value Added Per Capita

Figure 8: Expenditure Shares

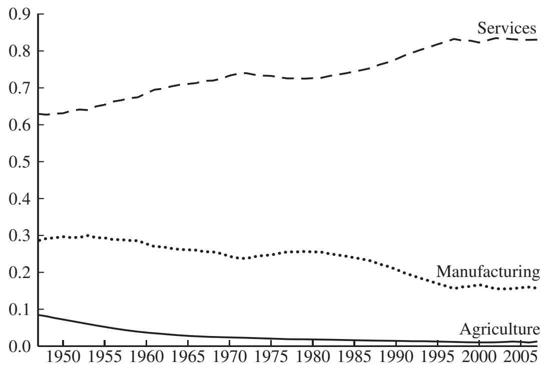


Figure 9: Price Indices (1947=1)

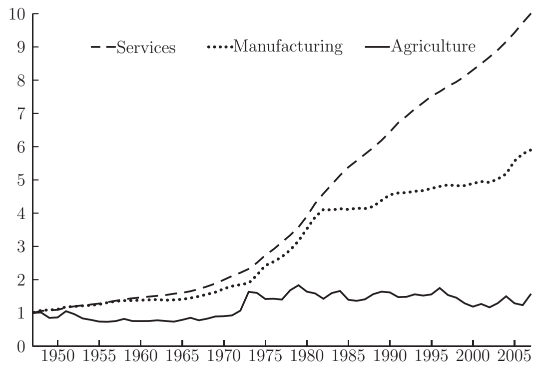
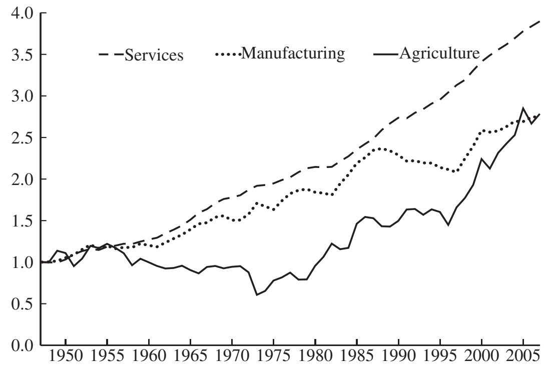


Figure 10: Quantity Indices (2000 chained dollars, 1947=1)



Although readily available, the data on value added and prices are not sufficient for our purposes. The reason is that value added data come from the production side of the national income and products accounts, and so they contain both consumption and investment. It is therefore necessary to devise a method to extract the consumption component from the production value added of each sector. This has not been sufficiently appreciated in the literature, which often proceeds by assuming that all investment is done in manufacturing.[20] This assumption is problematic, because since 1999 the BEA reports that the total value added in manufacturing has been consistently smaller than investment. We therefore need to properly extract the consumption component from the total value added in each sector. One contribution of our paper is to lay out a procedure that achieves this.

To carry out this extraction one needs to combine the value added data from the income side of the NIPA with the final expenditure data from the expenditure side of the NIPA. The complete details of this procedure are fairly involved, and so we relegate its description to Appendix C. Here we provide just a rough sketch of the procedure. A key difference between value added data from the income side and final expenditure data from the expenditure side is that the former are measured in what the BEA calls *producer's prices* whereas the latter are measured in *purchaser's prices*. From a practical perspective, the key difference is that purchaser's prices include distribution costs whereas producer's price do not (distribution costs are sales taxes and transport, wholesale, and retail services). For example, in the case of a shirt purchased from a retail outlet, the purchaser's price is the price paid by the consumer in the retail outlet whereas the producer's price is the price of the shirt when it leaves the factory.

In order to break final consumption expenditure into its value added components the first step therefore is to convert final consumption expenditure measured in purchaser's prices into those measured in producer's prices. This amounts to removing distribution costs from final consumption expenditure on goods and moving them into expenditure on services. Appendix C.1 explains the details of this calculation. Once this is done the second step is to use the input–output tables to determine the sectoral inputs in terms of value added that are required to deliver the final consumption expenditure. This involves an object called the total requirement

[20] Examples include Huffman and Wynne (1999) and Buera and Kaboski (2009).

matrix which is derived from the input–output tables. Appendix C.2 explains the details of this procedure.

Two points are worth stressing. First, since we are interested in the time series properties of consumption value added, and the structure of input–output relationships changes over time, an important feature of our calculation is that we use all annual input–output tables together with all benchmark tables that are available for the period 1947–2007. Second, when we break final consumption expenditure into its value added components we follow the BEA and treat imported goods as if they were produced domestically with the same technology that the U.S.A. uses to produce them. Given this assumption, we do not have to take a stand on whether intermediate goods are produced domestically or imported.[21]

Having broken final consumption expenditure into its value added components, we obtain data on consumption value added expenditure shares and chain–weighted prices and quantities, which are displayed in Figures 8–10. Note that these figures display the same qualitative pattern for consumption value added shares that we saw in the analogous figure for final consumption expenditure shares. Hence, both representations are consistent with the stylized facts about structural transformation. However, although the shares display similar qualitative behavior, there are some important differences in the behavior of relative prices and quantities. First, Figure 9 shows that while the price of services still increased the most, the price of manufacturing now increased by more than that of agriculture. Second, Figure 10 shows that relative quantities behave very differently. Whereas Figure 3 indicated substantial changes in relative quantities, Figure 10 suggests that the relative quantities of manufacturing and services now hardly change over the entire period, while the relative quantity of agriculture remains fairly constant after about 1970.

Given that relative prices changed substantially, the near constancy of relative quantities, particularly of manufacturing relative to services, suggests a very low degree of substitutability between the different components of consumption value added. Moreover, the near constancy of the relative agricultural quantity after 1970 suggests that nonhomotheticicities will not play as important a role as before. We will return to the significance of these observations below when

[21] Appendix C.2 explains this point in more detail.

Table 3: Results with Consumption Value Added

	(1)	(2)	(3)
σ	0.06	0	0
	(0.09)	–	–
$\bar{c}_a$	-121.92^{**}	-121.16^{**}	
	(9.28)	(14.25)	
$\bar{c}_s$	3474.89^{**}	3959.93^{**}	
	(693.05)	(363.96)	
ω_a	0.001	0.001	0.01^{**}
	(0.001)	(0.001)	(0.001)
ω_m	0.157^{**}	0.149^{**}	0.19^{**}
	(0.038)	(0.003)	(0.002)
ω_s	0.842^{**}	0.849^{**}	0.80^{**}
	(0.039)	(0.003)	(0.003)
AIC	−835.35	−843.89	−707.70
$RMSE_a$	0.005	0.005	0.010
$RMSE_m$	0.011	0.010	0.018
$RMSE_s$	0.011	0.011	0.024

Robust standard errors in parentheses * $p < 0.05$, ** $p < 0.01$
AIC is the Akaike information criterion, $RMSE_i$ is the root mean squared error for equation i.

we present the estimation results.

4.2 Results with Consumption Value Added

We follow the same procedure as was described previously in the context of estimating parameters using data on final consumption expenditure. Results are contained in Table 3. Column (1) reports the parameter estimates when we impose no restrictions. Strikingly, the point estimate of σ is equal to 0.06 and not statistically significantly different from zero, which in the absence of nonhomotheticities is basically the Leontief specification.[22] The nonhomothetic terms have the same signs as before. Column (2) shows the estimates when we impose $\sigma = 0$. Figure 11 shows that based on the estimates in Column (2), the fit of the model to the expenditure share data is again very good.

As before, it is of interest to ask how important income and relative price effects are in

[22] The corresponding Leontief utility function is given by $\min_{j=\{a,m,s\}}\{c_{jt}^{VA}/\phi_j\}$.

Table 4: Nonhomotheticity Terms Relative to Consumption Value Added from the Data

	1947	2007
$-p_a\bar{c}_a/C$	0.08	0.004
$p_s\bar{c}_s/C$	0.35	0.12
$-\bar{c}_a/c_a$	0.89	0.32
$\bar{c}_s/c_s$	0.56	0.14

accounting for the observed changes in the shares of consumption value added. One simple and revealing comparison about the relative importance of income effects for the two different data sets is to look at the size of the estimated $\bar{c}_i$'s relative to total consumption expenditure from the data. The first two rows of Table 4 shows that these ratios are now considerably smaller than in the case of final consumption expenditure, suggesting that income effects will be much less important. However, one needs to be somewhat careful with drawing immediate conclusions because the third row of the table shows that in 1947 the agricultural consumption value added from the data is now fairly close to $\bar{c}_a$. Intuitively, this can be understood by going back to Figure 10 above, which showed that the quantity of agricultural consumption value did not increase in the first part of our sample. Given that $\sigma = 0$, our estimated utility function delivers this by having agricultural consumption value added close to $\bar{c}_a$ initially.

To explore more systematically the importance of income versus relative price effects, Figure 12 shows the fit of the expenditure shares implied by the parameters of Column (2) when relative prices change as dictated by the data while total consumption expenditures are held constant at their 1947 value. While the fit deteriorates somewhat, it remains reasonably good. This suggests that relative price effects are the dominant force behind the changes in the expenditure shares of consumption value added. A second way of establishing this is to evaluate the ability of a homothetic specification to fit the data. To examine this, Column (3) in Table 3 presents the results under the restriction $\bar{c}_a = \bar{c}_m = \bar{c}_s = 0$. The first thing to note is that the Akaike information criterion increases somewhat, which suggests a worse goodness of fit. However, when we plot the expenditure share predicted by the homothetic specification in Figure 13, we can see that compared to the nonhomothetic specification of Figure 11, the fit remains reasonably good.

We conclude that the consumption value added data broadly support the homothetic preference specification used by Ngai and Pissarides, though in the somewhat extreme form of a Leontief specification, i.e. $\sigma = 0$. Since introspection would suggest substantial willingness to substitute across many commodities, some readers might question the empirical plausibility of preferences that do not permit any substitution across the consumption value added categories agriculture, manufacturing, and services. It is therefore important to understand exactly what the result $\sigma = 0$ means. Although having $\sigma = 0$ implies that there is no substitutability across the three categories, it is completely consistent with there being substantial substitution *within* each of these categories. In particular, since the categories are quite broad, having $\sigma = 0$ does not in any sense imply that there is no substitutability between all the different goods and services that individuals consume.

A simple example may be useful. Most readers will agree that there is some substitutability between the two activities of going to the movies and going to sporting events. When we represent these activities in consumption value added terms, we see that both of them involve some consumption of goods (e.g., the use of buildings) and some consumption of services (e.g., actors and athletes producing entertainment services). To us it seems reasonable to think that the key dimensions of substitution are within these two value added categories, i.e., that the key substitution is between the uses of buildings or the uses of athletes' and entertainers' time, rather than between goods and services per se. While this is not to suggest that one cannot think of specific examples with some substitution between specific goods and specific services, the key point we want to make is that there is likely to be considerably more substitutability within each of the value added categories.

Our conclusion that a low σ provides the best fit to value added data is related to independent research by Buera and Kaboski (2009). These authors ask whether there are parameters for which a canonical model of structural transformation can match the value added shares by sector between 1870 and 2000. They take as given initial GDP, a time series of overall TFP, and proxies for the price indices of the three sectors. They do not use information on investment, total consumption, and output from data, but let their model choose these quantities under the

Fit of Expenditure Shares with Consumption Value Added

Figure 11: Fit of Column (2)

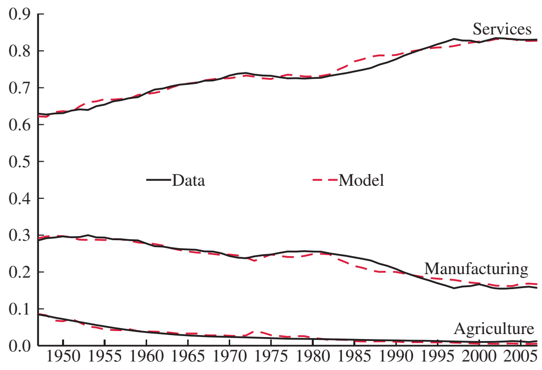


Figure 12: Fit of Column (2) with Income Fixed at 1947 Value

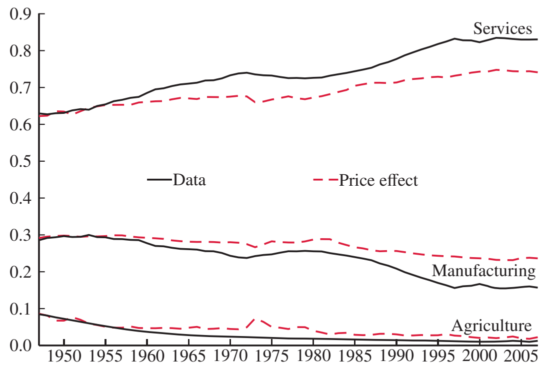
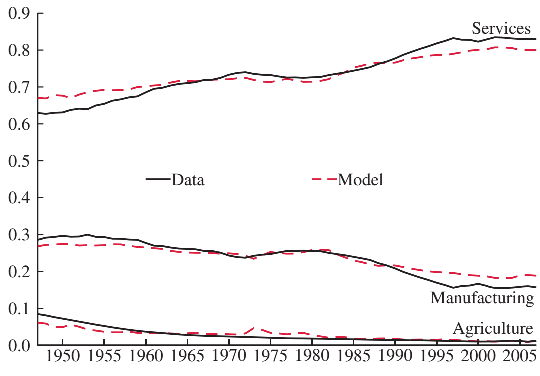


Figure 13: Fit of Homothetic Specification in Column (3)



assumption that all investment in the model is done in the manufacturing sector. Their preferred choice for σ comes out fairly low as well ($\sigma = 0.5$), but in contrast to us they conclude that their model cannot provide a good fit to the data.

This raises the question why the conclusions of the two studies are different. Upon closer inspection, there are several potentially important differences. To begin with, Buera–Kaboski consider data at ten–year intervals from 1870 to 2000, which cover a much larger range of expenditure shares than our data. Moreover, since our data are not available prior 1947, they use data for sector expenditure shares and prices that are not necessarily mutually consistent.[23] Lastly, Buera–Kaboski do not estimate the utility function given the expenditure shares of consumption value added, prices, and total consumption, but they determine it together with the sector production functions so as to match the shares of production value added.

To best evaluate what drives the differences between the conclusions of the two studies, it would be natural to redo our estimation exercise for 1947–2007 using the data of Buera–Kaboski. Unfortunately, this is not possible, since as noted earlier, investment has exceeded manufacturing value added since 1999, implying that we cannot extract the consumption part of production value added following their assumption that all investment is done in manufacturing. Independently of this, we do note that the series used by Buera–Kaboski to proxy for relative prices do display very different behavior than the true relative price series for value added over the period 1947–2007, suggesting that data differences probably play a significant role.

5 Discussion

5.1 Comparing the Results

Both estimation exercises that we carry out yield utility specifications that provide very good fits to their respective data sets. However, the specifications have very different properties: in one case it is close to the Stone–Geary specification used by Kongsamut et al. (2001) whereas in the other case it is close to the CES specification with low elasticity used by Ngai and Pis-

[23] Buera–Kaboski use the implicit deflator of services in NIPA and the producer price index of finished goods from the BLS. The former is based on gross sales while the latter is based on final expenditure. In contrast, we use price indices that are based on value added.

sarides (2007). The importance of relative price effects and income effects in accounting for changes in expenditure shares are therefore dramatically different in the two cases. In the case of final consumption expenditure income effects are the dominant force behind changes in the expenditure shares, whereas in the case of consumption value added relative price effects are the dominant force.

As we have stressed previously, given the technology for producing final expenditure categories from value added categories, there is an implicit mapping from preferences defined over final expenditure categories to preferences defined over value added categories. In this section we explore the properties of this mapping in order to reconcile the two very different estimated utility functions.

As before, the household has preferences over final consumption goods of the form given by equation (2). We assume that the household self–produces final consumption goods by combining the different consumption value added categories. We will derive the form of preferences over consumption value added that is implied by the preferences over final consumption goods and the production technology that specifies how the household obtains final consumption goods from consumption value added. Because our empirical strategy was to uncover preference parameters by estimating the expenditure systems, our approach will emphasize how the expenditure system for value added consumption is derived from the expenditure system for final expenditure consumption.

The first step in this derivation is to specify how final consumption goods are produced from the different value added categories. For expositional purposes it is convenient to assume that the corresponding production functions have the CES functional form:

$$c_{it}^{FE} = \left[\sum_{j\in\{a,m,s\}} (A_{it}\phi_{ji})^{\frac{1}{\eta_i}} \left(c_{jit}^{VA}\right)^{\frac{\eta_i-1}{\eta_i}} \right]^{\frac{\eta_i}{\eta_i-1}}, \tag{5}$$

where c_{jit}^{VA} is the value added from sector j that is used as an intermediate input in the production of the final consumption good c_{it}^{FE}, A_{it} determines the TFP of producing final consumption of category i, ϕ_{ji} are relative weights with $\sum_j \phi_{ji} = 1$, and $\eta_i > 0$ is the elasticity of substitution.

The household's demand functions for c_{jit}^{VA} are obtained by minimizing the costs of producing a given quantity c_{it}^{FE} subject to (5) while taking as given p_{jt}^{VA}. The demand functions take the familiar form:

$$p_{jt}^{VA} c_{jit}^{VA} = \frac{\phi_{ji} \left(p_{jt}^{VA}\right)^{1-\eta_i}}{\sum_{n\in\{a,m,s\}} \phi_{ni} \left(p_{nt}^{VA}\right)^{1-\eta_i}} p_{it}^{FE} c_{it}^{FE}, \tag{6}$$

where we have used the identity $\sum_{j\in\{a,m,s\}} p_{jt}^{VA} c_{jit}^{VA} = p_{it}^{FE} c_{it}^{FE}$.

The next step in the derivation of the demand system for consumption value added is to aggregate the demands for c_{jit}^{VA} to the demand for c_{jt}^{VA}. Summing equation (6) over i, we obtain:

$$p_{jt}^{VA} c_{jt}^{VA} = \sum_{i\in\{a,m,s\}} \frac{\phi_{ji} \left(p_{jt}^{VA}\right)^{1-\eta_i}}{\sum_{n\in\{a,m,s\}} \phi_{ni} \left(p_{nt}^{VA}\right)^{1-\eta_i}} p_{it}^{FE} c_{it}^{FE}. \tag{7}$$

One can express (7) as a standard demand system for consumption value added that depends only on the p_{jt}^{VA}'s and on $P_t C_t = \sum_{i\in\{a,m,s\}} p_{it}^{FE} c_{it}^{FE}$. This involves two steps; substitute in the demand functions for $p_{it}^{FE} c_{it}^{FE}$, which depend on p_{it}^{FE} and $P_t C_t$; use that final expenditure prices are given by the following price index:

$$p_{it}^{FE} = \left[\sum_{n\in\{a,m,s\}} A_{it} \phi_{ni} \left(p_{nt}^{VA}\right)^{1-\eta_i} \right]^{\frac{1}{1-\eta_i}}. \tag{8}$$

At a general level there is not that much that we can say about this resulting system. In general it will not even be consistent with preferences over value added consumptions that are of the form given by equation (2). Given our estimation results, the remainder of this subsection is devoted to the question of under what (if any) conditions the demand system (7) can be consistent with a Leontief utility function. To answer this question we note that if we assume:

$$\eta_i = 0 \quad \text{and} \quad \phi_{ji} = \phi_j \quad \forall i \in \{a, m, s\}, \tag{9}$$

then simple manipulation of (7) leads to:

$$p_{jt}^{VA} c_{jt}^{VA} = \frac{\phi_j p_{jt}^{VA}}{\sum_{n\in\{a,m,s\}} \phi_n p_{nt}^{VA}} \sum_{i\in\{a,m,s\}} p_{it}^{FE} c_{it}^{FE} = \frac{\phi_j p_{jt}^{VA}}{\sum_{n\in\{a,m,s\}} \phi_n p_{nt}^{VA}} P_t C_t. \tag{10}$$

This is readily seen to be the demand system that is implied by a Leontief utility function.

Table 5: Results for the Estimation of (7)

	Agriculture	Manufacturing	Services
η	0.01	0.002	0.01
	(0.01)	(0.01)	(0.01)
ϕ_{ai}	0.03**	0.01**	0.01**
	(0.00)	(0.00)	(0.002)
ϕ_{mi}	0.35**	0.34**	0.04**
	(0.01)	(0.05)	(0.01)
ϕ_{si}	0.62**	0.65**	0.95**
	(0.01)	(0.05)	(0.01)
AIC	−662.64	−584.63	−621.27

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$

The condition $\eta_i = 0$ means that the production functions (5) have the Leontief form and $\phi_{ji} = \phi_j$ means that the intermediate input from a given sector has the same weight in the production of all three final consumption goods. In this case, the aggregate intermediate inputs are not substitutable and each aggregate intermediate input has the same weight in the production of total final consumption as it has in the production of each of the three final consumption categories. Intuitively, this implies that the demand for intermediate inputs from a given sector is independent of the composition of final consumption.

Having isolated some theoretical conditions under which the value added demand system is consistent with Leontief preferences, we now turn to assessing the empirical relevance of these conditions. Given the observations of p_{jt}^{VA}, $p_{jt}^{VA}c_{jt}^{VA}$, and $p_{it}^{FE}c_{it}^{FE}$, we estimate the parameters η_i and ϕ_{ni} in equation (7) similar to the way that we estimated demand systems in the previous sections. The results are in Table (5). All of the point estimates for the η_i's are very small and none is statistically different from zero at the 5% level, implying that the first condition in (9) seems empirically reasonable. However, the second condition in (9) is statistically rejected by the data: while some values in a given row are very similar, others are fairly different. Nonetheless, it is still possible that the demand system generated by a Leontief utility function may provide a good fit to the data on consumption value added. To see why this is the case, note that (10) can be written as:

$$p_{jt}^{VA} c_{jt}^{VA} = \Phi_{jt} \frac{\phi_j p_{jt}^{VA}}{\sum_{n \in \{a,m,s\}} \phi_n p_{nt}^{VA}} P_t C_t, \tag{11}$$

where

$$\Phi_{jt} = \sum_{i \in \{a,m,s\}} \left(\frac{\phi_{ji} p_{jt}^{VA}}{\phi_j p_{jt}^{VA}} \frac{\sum_{n \in \{a,m,s\}} \phi_n p_{nt}^{VA}}{\sum_{n \in \{a,m,s\}} \phi_{ni} p_{nt}^{VA}} \right) \frac{p_{it}^{FE} c_{it}^{FE}}{P_t C_t}. \tag{12}$$

The demand system (11) is implied by the Leontief utility function if and only if $\Phi_{jt} = 1$. A sufficient condition for this to hold is that condition (9) holds. But more generally, as long as the Φ_{jt} are not too different from 1 the Leontief utility function may still provide a good fit to the data. It turns out that the departures from condition (9) are sufficiently small that the Leontief utility function provides a good fit.

5.2 Relative Merits of the Two Specifications

Each of the two specifications that we have considered has some relative merits over the alternative. In this subsection, we will discuss the merits in terms of applicability. We will not attempt to offer a view on which utility function is conceptually preferable. In the spirit of Lancaster (1966), each specification represents a different way of aggregating over a large set of characteristics that consumers value, and one can easily think of examples in which one or the other specification seems preferable.

In terms of applicability, one key issue is data availability. The relative advantage of the final consumption expenditure specification is that data on final consumption expenditure by category are readily available, not only from individual country sources but also in commonly used cross-country data sets such as the Penn World Table, which measure final consumption expenditure, as opposed to production. In contrast, consumption value added data are not readily available. To be sure, data on production value added by sector are readily available, but as we argued above, this is not the same as consumption value added.

When one thinks about integrating the consumer analysis into a general equilibrium setting that has a production side, then the relative merits are reversed. If, on the one hand, the arguments of the utility function are consumption value added, then one can include production in a consistent fashion by assuming value added production functions at the sector level. These, in turn, are easily connected to data since data on value added by sector are readily available. If, on the other hand, the arguments of the utility function are final consumption expenditure across categories of goods and services, then one either needs to write down a production structure that captures the complexities of the input–output relationships at the sector level, or one needs to find a representation of production that isolates the contribution of capital and labor to the production of final expenditure categories. While this can be done, it is more difficult than working directly with sectoral value added production functions.[24]

Additionally, to the extent that one desires utility functions that are consistent with aggregation, given our estimates, the consumption value added specification is preferable. This is due to the fact that the nonhomotheticity terms are relatively unimportant in this case. As a result, the utility specification for consumption value added aggregates for a larger set consumption expenditure than the utility specification for final consumption expenditure.[25] Moreover, the homothetic specification from Column (2) of Table 3 still provided a very good fit in the case of consumption value added, and in this case aggregation always holds.

5.3 Additional Measurement Issues

In this subsection we note five measurement issues that are relevant not only for our analysis but for virtually any study of this sort. The first issue is that government services are often badly measured (e.g., because value added is "approximated" by the corresponding wage bill). One may therefore wonder to what extent our estimation results are driven by the behavior of badly measured government services. To address this point, we have redone all of the estimations without government services.[26] Naturally, this reduces the quantity of services consumed, and so it lowers the estimates of the relative weight on services and the nonhomotheticity term $\bar{c}_s$. But the important question is what happens to the estimates of the elasticity of substitution and the nonhomotheticity term for agriculture. Using consumption value added data, we find that

[24]See Valentinyi and Herrendorf (2008).

[25]Formally, condition (13) in Appendix A holds for larger set of C_n.

[26]Our initial results implicitly assumed that households were purchasing government services at the price p_s. In contrast, here we remove government services and implicitly assume that whatever utility individuals obtain from these services does not affect the marginal rate of substitution between categories of private consumption.

Table 6: Decomposition of Increase in Expenditure Share of Services Consumption Value Added (accumulated 1947–2007)

Category	**%**
Finance, Insurance, Real Estate, Rental, and Leasing	48.8
Professional and Business Services	41.5
Health Care and Social Assistance	26.3
Information	8.3
Utilities	3.4
Educational Services	2.9
Government	1.0
Arts, Entertainment, Recreation, Accommodation, Food Services, and Other	−0.5
Trade and Transport	−31.7
	100

these results are literally unchanged, while using final consumption expenditure data the results are affected only somewhat.[27] We conclude from this exercise that our results are not chiefly driven by the behavior of badly measured government services.

Second, and related, an important issue when examining time series changes in prices and quantities is the extent to which the data take proper account of quality improvements. Failure to do so will bias the decomposition of expenditure into price and quantity components. A key limitation of the official data that we have used in our analysis is that effectively no corrections are made to allow for quality improvements in services. While this is a common problem that we cannot do anything about, we think that it is worth keeping in mind.

Third, consumption of durable goods typically does not equal expenditure on durable goods. For housing, which is by far the most prominent example of durables, the BEA takes account of this and imputes the rents for owner–occupied houses. For all other durables, the BEA reports expenditure (or value added) only, which forces us to associate the expenditures on these durables with current consumption. This implies, for example, that current period utility from automobiles is derived solely from current period sales (or production) of automobiles, and so we do not attribute any current period utility flow to the stock of automobiles purchased in previous periods. Because we are focused on longer term trends in aggregate data, this is not

[27] The precise results can be found in Appendix D.

likely to be as serious as it would be in looking at individual data, or business cycle fluctuations; but it is an issue worth noting.

Fourth, our model has abstracted from home production. Aguiar and Hurst (2007) and Ramey and Francis (2009) both documented a sharp drop in time devoted to home production associated with the dramatic increase in the participation rate of married women. To the extent that this has led to a substitution away from home produced services toward market produced services, our data may reflect an upward bias in the extent of the increase in the expenditure share of services.

Fifth, and related, one issue with sectoral data is the possibility that reallocation of resources across sectors reflects a relabeling of activity due to outsourcing, as opposed to fundamental shifts of economic activity across sectors. For example, if a car manufacturer changes from having in–house security guards at its establishments to purchasing security services from an outside firm, the data will record this as a movement of value added across sectors.[28] This phenomenon will bias the measurement of changes in the expenditure shares of consumption value added. However, this bias is not likely to be a major driving force of structural transformation. The main reason is that industry classifications are done at the establishment level, implying that all in–house services provided at a central administrative office (headquarters) or a separate service–providing unit are always classified as service industries.

There are two additional ways of establishing that outsourcing is not the major force behind structural transformation. First, Table 6 decomposes the accumulated increase in the expenditure share of service consumption value added into the contributions of ten subcategories of services, where outsourced services are part of the subcategory Professional and Business Services. Although this category is the second biggest contributor to the overall increase in the expenditure share of services, more than half of the increase is accounted for by other categories. Moreover, it is reasonable to think that a substantial share of the increase in business and professional services reflects purchases directly made by consumers, in which case they would not be subject to outsourcing. A second way of establishing that outsourcing is not the major force behind structural transformation is to look at what happened to final consumption

[28]Fuchs (1968) suggested that this is one of the driving forces behind the process of structural transformation.

expenditure, instead of consumption value added, as final consumption expenditure are not affected by outsourcing. To stay with the example of the car manufacturer, all that matters with final consumption expenditure is how much is spent on purchases of cars. Holding the price and quantity of security services fixed, it does not matter if the security services that are implicitly reflected in the price of cars were supplied in–house or outsourced. The fact that the changes in the shares are very evident in the final consumption expenditure data confirms that the process of structural transformation is not mainly a process of outsourcing.

6 Conclusion

What utility function should one use in applied work on structural transformation and related issues? This paper sought to provide an answer to this simple question by examining the behavior of household expenditure shares for the US economy over the period 1947 to 2007. The first contribution of this paper is to clarify that given common practice in specifying multi-sector general equilibrium models, this question requires two answers, one each for two different methods of defining commodities in such models.

The second contribution of this paper is to supply the two answers. A priori there is little guidance as to how different (or similar) the two answers might be. It is very noteworthy that we find the answers to be dramatically different in terms of their basic properties. Interestingly, each of the answers takes a very simple form. If one adopts the final consumption expenditure specification, then a Stone-Geary utility function provides a very good fit to the US time series data. If instead one adopts the consumption value added specification, then a homothetic specification of the Leontief type provides a very good fit to the data.

A third contribution of the paper, which is a necessary intermediate input into the estimation of the utility function based on consumption value added, was to develop and execute a procedure for producing time series data on consumption value added. This requires extracting the component of total value added by sector that corresponds to consumption value added.

While the utility functions that we estimate are specifically relevant for models of structural transformation, some of the basic messages of the analysis are much more general. Specifically, researchers must be careful to apply consistent definitions of commodities on both the household and production sides when connecting models with data in any multi–sector general equilibrium analysis. Changing the true definition of what is meant, for example, by the label "services" has implications not only on the household side for what form of utility function is appropriate, but also on the production side for such things as the measurement of productivity growth. This has important implications for comparing results across studies and for the practice of importing parameter values across studies.

There are several dimensions along which it will be important to extend the analysis carried out here. For example, in this paper we have only analyzed the evolution of expenditure shares and prices in one country – the postwar U.S. It is also of interest to extend this analysis to a larger set of countries, in particular to situations which feature a larger range of real incomes. This will be useful in assessing the extent to which one can account for the process of structural transformation with stable preferences.

References

Acemoglu, Daron and Veronica Guerrieri, "Capital Deepening and Non–Balanced Economic Growth," *Journal of Political Economy*, 2008, *116*, 467–498.

Aguiar, Mark and Eric Hurst, "Measuring Leisure: The Allocation of Time Over Five Decades," *Quarterly Journal of Economics*, 2007, *122*, 969–1006.

Atkeson, Andrew and Masao Ogaki, "Wealth–varying Intertemporal Elasticity of Substitution: Evidence from Panel and Aggregate Data," *Journal of Monetary Economics*, 1996, *38*, 507–534.

Bah, El-Hadj, "Structural Transformation in Developed and Developing Countries," mamuscript, Arizona State University., Tempe, AZ 2008.

Barnett, William A. and Apostolos Serletis, "Consumer Preferences and Demand Systems," *Journal of Econometrics*, 2008, *147*, 210–224.

Baumol, William J., "Macroeconomics of Unbalanced Growth: The Anatomy of the Urban Crisis," *American Economic Review*, 1967, *57*, 415–426.

Blundell, Richard, "Consumer Behaviour: Theory and Evidence – A Survey," *Economic Journal*, 1988, *98*, 16–65.

Buera, Francisco J. and Joseph P. Kaboski, "Can Traditional Theories of Structural Change Fit The Data?," *Journal of the European Economic Association*, 2009, *7*, 469–477.

Bureau of Economic Analysis, "Concepts and Methods of the U.S. Input–Output Accounts," 2006. http://bea.gov/bea/papers/IOmanual_092906.pdf.

Caselli, Francesco and Wilbur John Coleman, "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 2001, *109*, 584–616.

Deaton, Angus and John Muellbauer, "An Almost Ideal Demand System," *American Economic Review*, 1980, *70*, 312–326.

Duarte, Margarida and Diego Restuccia, "The Role of the Structural Transformation in Aggregate Productivity," *Quarterly Journal of Economics*, 2010, *125*, 129–173.

Echevarria, Cristina, "Changes in Sectoral Composition Associated with Economic Growth," *International Economic Review*, 1997, *38*, 431–452.

Foellmi, Reto and Josef Zweimüller, "Structural Change, Engel's Consumption Cycles, and Kaldor's Facts of Economic Growth," *Journal of Monetary Economics*, 2008, *55*, 1317–1328.

Fuchs, Victor, *The Service Economy*, New York: Columbia University Press, 1968.

Geary, Roy C., "A Note on A Constant–Utility Index of the Cost of Living," *Review of Economic Studies*, 1950-1951, *18*, 65–66.

Gollin, Douglas, Stephen L. Parente, and Richard Rogerson, "The Role of Agriculture in Development," *American Economic Review, Papers and Proceedings*, 2002, *92*, 160–164.

Herrendorf, Berthold, James Schmitz Jr, and Arilton Teixeira, "Transportation and Development: Insights from the US, 1840–1860," *Forthcoming: International Economic Review*, 2011.

_ **, Richard Rogerson, and Ákos Valentinyi**, "Growth and Structural Transformation," Chapter prepared for the Handbook of Economic Growth 2011.

Houthakker, Hendrik S. and Lester Taylor, *Consumer Demand in the United States: Analysis and Projections*, Cambridge, MA: Harvard University Press, 1970.

Huffman, Gregory W. and Mark A. Wynne, "The Role of Intratemporal Adjustment Costs in a Multisector Economy," *Journal of Monetary Economics*, 1999, *43*, 317–350.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie, "Beyond Balanced Growth," *Review of Economic Studies*, 2001, *68*, 869–882.

Kuznets, Simon, *Modern Economic Growth*, New Haven: Yale University Press, 1966.

Laitner, John, "Structural Change and Economic Growth," *Review of Economic Studies*, 2000, *67*, 545–561.

Lancaster, Kelvin, "A New Approach to Consumer Theory," *Journal of Political Economy*, 1966, *74*, 132–157.

Landefeld, J. Steven and Robert P. Parker, "BEA's Chain Indexes, Time Series, and Measures of Long Term Economic Growth," in "Survey of Current Business," Bureau of Economic Analysis, May 1997.

Messina, Julian, "The Role of Product Market Regulations in the Process of Structural Change," *European Economic Review*, 2006, *50*, 1863–1890.

Ngai, L. Rachel and Chrisopher A. Pissarides, "Structural Change in a Multisector Model of Growth," *American Economic Review*, 2007, *97*, 429–443.

_ and _ , "Trends in Hours and Economic Growth," *Review of Economic Dynamics*, 2008, *11*, 239–256.

Pollak, Robert and Terence J. Wales, "Estimation of the Linear Expenditure System," *Econometrica*, 1969, *37*, 611–628.

Ramey, Valery A. and Neville Francis, "A Century of Work and Leisure," *AEJ: Macro*, 2009, *1*, 189–224.

Rogerson, Richard, "Structural Transformation and the Deterioration of European Labor Market Outcomes," *Journal of Political Economy*, 2008, *226*, 235–259.

Sato, Ryuzo, "The Most General Class of CES Functions," *Econometrica*, 1975, *43*, 999–1003.

Stone, Richard, "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," *Economic Journal*, 1954, *64*, 511–527.

ten Raa, Thijs, *The Economics of Input-Output Analysis*, Cambridge, UK: Cambridge University Press, 2005.

Valentinyi, Akos and Berthold Herrendorf, "Measuring Factor Income Shares at the Sectoral Level," *Review of Economic Dynamics*, 2008, *11*, 820–835.

Appendix A: Aggregation of Demand Functions

Consider N households indexed by $n = 1, ..., N$. Each household solves:

$$\max_{c_a^n, c_m^n, c_s^n} \left[\omega_a^{\frac{1}{\sigma}} (c_a^n + \bar{c}_a)^{\frac{\sigma-1}{\sigma}} + \omega_m^{\frac{1}{\sigma}} (c_m^n + \bar{c}_m)^{\frac{\sigma-1}{\sigma}} + \omega_s^{\frac{1}{\sigma}} (c_s^n + \bar{c}_s)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

$$\text{s.t.} \quad p_a c_a^n + p_m c_m^n + p_s c_s^n \leq C_n.$$

Let the parameters and the income distribution be such that for all $n \in \{1, ..., N\}$ household expenditure exceed a minimum level:

$$C_n > \sum_{i=a,m,s} p_i \max\{-\bar{c}_i, 0\}. \tag{13}$$

Then the solution to each household's problem is interior and the first–order conditions are

$$\left(\frac{\omega_a}{\omega_s}\right)^{\frac{1}{\sigma}} \left(\frac{c_a^n + \bar{c}_a}{c_s^n + \bar{c}_s}\right)^{-\frac{1}{\sigma}} = \frac{p_a}{p_s},$$
$$\left(\frac{\omega_m}{\omega_s}\right)^{\frac{1}{\sigma}} \left(\frac{c_m + \bar{c}_m}{c_s^n + \bar{c}_s}\right)^{-\frac{1}{\sigma}} = \frac{p_m}{p_s},$$

which can be rewritten as

$$\frac{p_a}{p_s} \frac{c_a^n + \bar{c}_a}{c_s^n + \bar{c}_s} = \frac{\omega_a}{\omega_s} \left(\frac{p_a}{p_s}\right)^{1-\sigma},$$
$$\frac{p_m}{p_s} \frac{c_m^n + \bar{c}_m}{c_s^n + \bar{c}_s} = \frac{\omega_m}{\omega_s} \left(\frac{p_m}{p_s}\right)^{1-\sigma}.$$

This gives the demand functions

$$p_a(c_a^n + \bar{c}_a) = \frac{p_a(c_a^n + \bar{c}_a) + p_m(c_m^n + \bar{c}_m) + p_s(c_s^n + \bar{c}_s)}{1 + \dfrac{\omega_m}{\omega_a}\left(\dfrac{p_m}{p_a}\right)^{1-\sigma} + \dfrac{\omega_s}{\omega_a}\left(\dfrac{p_s}{p_a}\right)^{1-\sigma}},$$

$$p_m(c_m^n + \bar{c}_m) = \frac{p_a(c_a^n + \bar{c}_a) + p_m(c_m^n + \bar{c}_m) + p_s(c_s^n + \bar{c}_s)}{1 + \dfrac{\omega_a}{\omega_m}\left(\dfrac{p_a}{p_m}\right)^{1-\sigma} + \dfrac{\omega_s}{\omega_m}\left(\dfrac{p_s}{p_m}\right)^{1-\sigma}},$$

$$p_s(c_s^n + \bar{c}_s) = \frac{p_a(c_a^n + \bar{c}_a) + p_m(c_m^n + \bar{c}_m) + p_s(c_s^n + \bar{c}_s)}{1 + \dfrac{\omega_a}{\omega_s}\left(\dfrac{p_a}{p_s}\right)^{1-\sigma} + \dfrac{\omega_m}{\omega_s}\left(\dfrac{p_i}{p_s}\right)^{1-\sigma}}.$$

Adding up over all households, we obtain:

$$p_a(c_a + N\bar{c}_a) = \frac{p_a(c_a + N\bar{c}_a) + p_m(c_m + N\bar{c}_m) + p_s(c_s + N\bar{c}_s)}{1 + \dfrac{\omega_m}{\omega_a}\left(\dfrac{p_m}{p_a}\right)^{1-\sigma} + \dfrac{\omega_s}{\omega_a}\left(\dfrac{p_s}{p_a}\right)^{1-\sigma}},$$

$$p_m(c_m + N\bar{c}_m) = \frac{p_a(c_a + N\bar{c}_a) + p_m(c_m + N\bar{c}_m) + p_s(c_s + N\bar{c}_s)}{1 + \dfrac{\omega_a}{\omega_m}\left(\dfrac{p_a}{p_m}\right)^{1-\sigma} + \dfrac{\omega_s}{\omega_m}\left(\dfrac{p_s}{p_m}\right)^{1-\sigma}},$$

$$p_s(c_s + N\bar{c}_s) = \frac{p_a(c_a + N\bar{c}_a) + p_m(c_m + N\bar{c}_m) + p_s(c_s + N\bar{c}_s)}{1 + \dfrac{\omega_a}{\omega_s}\left(\dfrac{p_a}{p_s}\right)^{1-\sigma} + \dfrac{\omega_m}{\omega_s}\left(\dfrac{p_i}{p_s}\right)^{1-\sigma}},$$

where

$$c_i \equiv \sum_{n=1}^{N} c_i^n.$$

Let $C \equiv \sum_{n=1}^{N} C^n$. If the stand–in household solves

$$\max_{c_a, c_m, c_s} \left[\omega_a(c_a + N\bar{c}_a)^{\frac{\sigma-1}{\sigma}} + \omega_m(c_m + N\bar{c}_m)^{\frac{\sigma-1}{\sigma}} + \omega_s(c_s + N\bar{c}_s)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

$$\text{s.t.} \quad p_a c_a + p_m c_m + p_s c_s \leq C,$$

then its choices satisfy

$$c_i = \sum_{n=1}^{N} c_i^n.$$

In other words, there is aggregation.

Appendix B: Data Sources and Sector Assignment

B.1: Data Sources

All data are for the U.S. during 1947–2007.

We calculate a per capita quantity by dividing the total quantity by the population size. We take the population size from Table 7.1: "Selected Per Capita Product and Income Series in Current and Chained Dollars".

The construction of final consumption expenditure data is based on standard NIPA tables from the BEA. We use the most recent NIPA data released in August 2009 which incorporates the last comprehensive revision. In particular, we use data from the following tables:

- Table 2.4.3: "Real Personal Consumption Expenditures by Type of Product, Quantity Indexes"; Table 2.4.5: "Personal Consumption Expenditures by Type of Product".

- Table 3.10.3: "Real Government Consumption Expenditures and General Government Gross Output, Quantity Indexes"; Table 3.10.5: "Government Consumption Expenditures and General Government Gross Output"

The construction of production value added data by sector is based on the Annual Industry Accounts, which contain current dollar value added and quantity indices by industry based on chain weighted methods. The value added by industry data is consistent with the NAICS for the entire period 1947–2007.[29]

The construction of consumption value added (as opposed to production value added) is based on two main data sources: the annual expenditure data described above and the total

[29] `http://www.bea.gov/industry/gpotables/AllTables.zip`

requirement matrices from the IO Tables. In the next subsection, we describe in detail how these two data sources are combined to obtain consumption value added. Here we just describe the exact data sources. There are benchmark IO Tables and annual IO Tables. Benchmark IO Tables are available for 1947, 1958, 1963, 1967, 1972, 1977, 1982, 1987, 1992, 1997 and 2002.[30] Annual IO Tables are available for each year during the period 1998–2007.[31] An important additional data source are the so called "Bridge Tables for Personal Consumption Expenditure", which are available for the 1997 and 2002 benchmark IO Tables. Bridge Tables link IO Tables with the standard expenditure data of the BEA. In particular, they report how personal consumption expenditure in the IO Tables are related to those in the BEA expenditure tables. If we don't have IO Tables for a particular year, then we use linear interpolation between the years for which IO Tables are available.

B.2: Sector Assignment

When we use final consumption expenditure data, the three sectors contain the following BEA *commodities*:

- Agriculture: "food and beverages purchased for off-premises consumption"
- Manufacturing: "durable goods"; "nondurable goods" excluding "food and beverages purchased for off-premises consumption"
- Services: "services"; "government consumption expenditure"

When we use value added data, the three sectors contain the following BEA *industries*:

- Agriculture: "farms"; "forestry, fishing, and related activities"
- Manufacturing: "construction"; "manufacturing"; "mining"
- Services: all other industries including "government industries"

[30] http://www.bea.gov/industry/io_benchmark.htm
[31] http://www.bea.gov/industry/io_annual.htm

Appendix C: Calculating Consumption Value Added

C.1: Constructing Final Expenditure in Producer's Prices

C.1.1: Disaggregation to six sectors

To obtain final consumption expenditure in producer's prices from the available data on final consumption expenditure in purchaser's prices, we need to remove the distribution costs from the different goods categories and move them to services. For two reasons, this requires further disaggregation. First, we calculate the distribution costs for retail, wholesale and transportation services from the expenditure on the sector Trade and Transport. We therefore, need to separate Trade and Transport from the rest of services. Second, the expenditure on mining involve distribution costs whereas those on construction do not, so we need to separate the two from other manufacturing. We therefore consider the following six sectors: Agriculture, Mining, Construction, Manufacturing, Trade and Transport, and Services excluding Trade and Transport, which we index by $i \in \{Ag, Mi, Co, Ma, TT, Se\}$, which aggregate to our model sectors in the obvious way: $a = \{Ag\}$, $m = \{Mi, Co, Ma\}$, $s = \{TT, Se\}$. Note that while we use the BEA classification for Agriculture, Mining, Construction, and Manufacturing , the sector Trade and Transport combines "Wholesale Trade", "Retail Trade" and "Transportation and Warehousing".

We should mention a potential problem that arises from the reclassification of industries over time. In particular, while the BEA now publishes GDP by industry data based on the NAICS for the whole period 1947–2007, it still publishes the underlying input–output tables for the subperiod 1947–1977 based on the different SIC's. Fortunately, many of the reclassifications from the SIC's to the NAICS happened at finer levels of disaggregation than we study here, and so they do not affect the aggregates of the six sectors we have just introduced. However, there are some exceptions that we need to reclassify to make the input–output tables consistent with the GDP by industry data. The most important example is the "Publishing Industries", which the SIC has as manufacturing industry and the NAICS has as a service industry.

C.1.2: Removing distribution costs from personal consumption expenditure

We now explain how to remove distribution costs from personal consumption expenditure.

The expenditure side of GDP values personal consumption expenditure at purchaser's prices and it disaggregates them into the expenditure on goods, trade and transportation, and services excluding trade and transportation:

$$PC^{Pu} = PC^{Pu}_{Gs} + PC^{Pu}_{TT} + PC^{Pu}_{Se}.$$

Goods consist of "durable and nondurable goods" excluding "food and beverages purchased for off-premises consumption", trade and transportation consists of "public transportation", and services consist of "services" excluding "public transportation".

We start by removing distribution costs from personal consumption expenditure on goods. To go from purchaser's to producer's prices, we calculate the distribution margins $DM_{PC_{Gs}}$ by using the fact that in the IO Tables personal consumption expenditure on trade and transportation consists of all transportation expenditure whereas PC^{Pu}_{TT} consists only of "public transportation" that households explicitly purchase. Hence, the difference between the two equals the distribution costs of goods that household purchase indirectly when purchasing goods, and so:

$$DM_{PC_{Gs}} = \frac{(PC^{IO}_{TT} - PC^{Pu}_{TT})}{(PC^{IO}_{TT} - PC^{Pu}_{TT}) + (PC^{IO}_{Ag} + PC^{IO}_{Mi} + PC^{IO}_{Co} + PC^{IO}_{Ma})},$$
$$PC^{Pr}_{Gs} = (1 - DM_{PC_{Gs}})PC^{Pu}_{Gs}.$$

We continue by removing distribution costs from personal consumption expenditure on services. This is straightforward because the IO Tables suggest that personal consumption expenditure on services involve negligible distribution costs. Therefore:

$$PC^{Pr}_{Se} = PC^{Pu}_{Se}.$$

Given that we have calculated PC^{Pr}_{Gs}, we now disaggregate it into the components PC^{Pr}_{Ag}, PC^{Pr}_{Mi}, PC^{Pr}_{Co} and PC^{Pr}_{Ma}. The IO Tables report that PC^{IO}_{Mi} are very small and that PC^{IO}_{Co} are zero in all years. We therefore set $PC^{Pr}_{Mi} = PC^{Pr}_{Co} = 0$. This leaves us with the task of splitting PC^{Pr}_{Gs} between PC^{Pr}_{Ag} and PC^{Pr}_{Ma}. First, we calculate expenditures on food at producer prices, PC^{Pr}_{Food}.

Expenditure on food is "food and beverages purchased for off-premises consumption". We remove distribution costs by applying the distribution margin of goods that we calculated above, $PC^{Pr}_{Food} = (1 - DM^{PC}_{Gs})PC^{Pu}_{Food}$. Next, since PC^{Pr}_{Food} contains both unprocessed and processed food, we need to take processed food out to obtain the expenditure on agricultural commodities. We use that PC^{IO}_{Ag} are the expenditure on agricultural goods without processed food. Defining $\Phi_1 \equiv PC^{IO}_{Ag}/PC^{Pr}_{Food}$, we have

$$PC^{Pr}_{Ag} = \Phi_1 PC^{Pr}_{Food}.$$

In sum, the components of personal consumption expenditure in producer's prices are obtained as follows:

$$\begin{aligned}
PC^{Pr}_{Ag} &= \Phi_1 PC^{Pr}_{Food} \\
PC^{Pr}_{Mi} &= 0, \\
PC^{Pr}_{Co} &= 0, \\
PC^{Pr}_{Ma} &= (1 - DM_{PC_{Gs}})PC^{Pu}_{Gs} - PC^{Pr}_{Ag}, \\
PC^{Pr}_{TT} &= PC^{Pu}_{TT} + DM_{PC_{Gs}} PC^{Pu}_{Gs}, \\
PC^{Pr}_{Se} &= PC^{Pu}_{Se}.
\end{aligned}$$

C.1.3: Removing distribution costs from government consumption expenditure

We now explain how to remove distribution costs from final expenditure on government consumption.

In the IO Tables, the general government appears as a production industry and as a commodity. In the expenditure side of GDP, government consumption expenditure at purchaser's prices are defined as the gross output of the general government industry minus own account investment and sales to other sectors.

The treatment of the gross output of the general government industry changed in 1998. Before 1998, it was defined as its value added GC^{Pu}_{VA} (compensation of general government employees plus consumption of general government fixed capital). All intermediate inputs were consequently treated as final government expenditure on these goods. Since 1998, the gross output of the general government industry has included intermediate goods, that is, it is defined as the sum of value added GC^{Pu}_{VA}, purchased intermediate goods, GC^{Pu}_{Gs}, and purchased intermediate services, GC^{Pu}_{Se}.

We start with the period 1947–1997. During this period, the IO Tables show that GC^{IO}_{Ag} and GC^{IO}_{Mi} are small, so we set $GC^{Pr}_{Ag} = GC^{Pr}_{Mi} = 0$. The distribution margins of government consumption expenditure in the 1997 IO Tables on average equal 18% of the distribution margins of personal consumption expenditure, so we set $DM_{GC_{Gs}} = 0.18 \cdot DM_{PC_{Gs}}$. Next we calculate GC^{Pr}_{Co}. The raw IO Tables distinguish between government expenditure on "maintenance and repair construction" and on "new construction". First, we calculate

$$\Phi_2 \equiv \frac{\text{government expenditure on maintenance and repair construction}}{\text{depreciation on government structures}},$$

where the depreciation on government structures is taken from Table 7.3: "Current-Cost Depreciation of Government Fixed Assets" of the BEA Fixed Assets Tables. We then calculate GC^{Pr}_{Co} by multiplying Φ_2 with depreciation on government structures.

In sum, for the period 1947–1997, we calculate the variables of interest as:

$$GC^{Pr}_{Ag} = 0, \tag{14a}$$

$$GC^{Pr}_{Mi} = 0, \tag{14b}$$

$$GC^{Pr}_{Co} = \Phi_2 \text{Depreciation on government structures}, \tag{14c}$$

$$GC^{Pr}_{Ma} = (1 - DM_{GC_{Gs}})GC^{Pu}_{Gs} - GC^{Pr}_{Co}, \tag{14d}$$

$$GC^{Pr}_{TT} = DM_{GC_{Gs}} GC^{Pu}_{Gs}, \tag{14e}$$

$$GC^{Pr}_{Se} = GC^{Pu}_{VA} + GC^{Pu}_{Se} - \text{Sales to other sectors} - \text{Own account investment}. \tag{14f}$$

The rationale behind the fourth equation is that own account investment typically involves goods, so it has to be taken out of government intermediate good expenditure. The fifth equation is just an implication of the fourth equation. The last equation expresses that government expenditure on services are equal to the value added representing the service flow from government capital and employees plus the services purchased as intermediate input net of what is invested on own account sold to other sectors, which typically are general government services. The equation reflects that typically services do not have distribution costs, so services evaluated at producer's and purchaser's prices are the same.

For the period 1998–2007, government consumption expenditure in the IO Tables almost exclusively consist of expenditure on general government services. Since services have no distribution costs, we set:

$$GC^{Pr}_{Ag} = GC^{Pr}_{Mi} = GC^{Pr}_{Co} = GC^{Pr}_{Ma} = GC^{Pr}_{TT} = 0,$$
$$GC^{Pr} = GC^{Pu}.$$

C.2: Linking Consumption Expenditures to Value Added

The total requirement matrix (henceforth TR Matrix) links the income and the expenditure side of GDP. We now explain how to use the TR Matrix to obtain the value added in producer's prices that are generated by the final expenditure on consumption in producer's prices, which we have just constructed in the previous subsection. We use the language and the notation of the BEA to the extent possible. For further explanation see ten Raa (2005) and Bureau of Economic Analysis (2006).

The way in which the TR Matrix is calculated changed in 1972. So for years prior to 1972, the IO Tables assumed that each industry produces one commodity and that each commodity is produced in exactly one industry. For years after 1972 the IO Tables have taken account of the fact that industries can produce more than one commodity and that the same commodity can be produced in different industries.

We start by explaining the TR Matrix prior to 1972. We denote the number of industries by n, which before 1972 equals the number of commodities. Domestically produced commodities are purchased either by domestic industries (intermediate expenditure) or by final users (final uses or final expenditure). Final uses include both domestic final uses and exports, where exports can be either intermediate or final foreign uses. Domestic industries produce gross output and the difference between gross output and intermediate expenditure is industry value added.

Let $\boldsymbol{A}$ denote the $(n \times n)$ transaction matrix.[32] Rows are associated with commodities and columns with industries: entry ij shows the dollar amount of commodity i that industries j uses per dollar of output it produces. Note that these commodities may have been produced domestically or imported. Let $\boldsymbol{q}$ denote the $(n \times 1)$ output vector of domestically produced commodities. Element i records the sum of the dollar amounts of commodity i that are delivered to other domestic industries as intermediate inputs and to final uses. Let $\boldsymbol{g}$ denote the $(n \times 1)$ industry output vector. Element j records the dollar amount of output of industry j. Let $\boldsymbol{e}$ denote the $(n \times 1)$ vector of expenditures on final uses. Element i records the dollar amount of final uses of the domestically produced commodity i, so component e_i reports domestic private and public consumption, domestic investment, and net exports of commodity i.

Two identities link these vectors and with the TR Matrix:

$$\boldsymbol{q} = \boldsymbol{A}\boldsymbol{g} + \boldsymbol{e}, \tag{15}$$

$$\boldsymbol{q} = \boldsymbol{g}. \tag{16}$$

The first identity states that the dollar amount of domestically produced output of each commodity equals the sum of intermediate uses plus the final uses of that commodity. The second identity states that total value of output of industry i equals to the total value of commodity i, which is trivially true here because each industry is assumed to produce one distinct commodity. We can solve these two equations for g:

$$\boldsymbol{g} = (\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{e}, \tag{17}$$

where $\boldsymbol{I}$ is the $(n \times n)$ identity matrix (1 in the diagonal and zero elsewhere). $\boldsymbol{R} \equiv (\boldsymbol{I} - \boldsymbol{A})^{-1}$ is called the total requirements matrix. Rows are associated with industries and columns with commodities. Entry ji shows the dollar value of industry j's production that is required, both directly and indirectly, to deliver one dollar of the domestically produced commodity i to final uses including net exports.

[32] Matrices and vectors are in bold symbol throughout the paper.

We continue by explaining the TR Matrix after 1972, so now the IO Tables take account of the fact that an industry may produce more then one commodity and that a commodity may be produced in different industries. In general, the number of industries n will then differ from the number of commodities. We call the number of commodities m. This implies that we no longer have one transaction matrix, but a use and a make matrix. $\boldsymbol{B}$ denotes the $(m \times n)$ use matrix. Entry ij shows the dollar amount of commodity i that industries j uses per dollar of output it produces. Again note that these commodities may have been produced domestically or imported. $\boldsymbol{W}$ denotes the $(n \times m)$ make matrix. Rows are associated with industries and columns with commodities: entry ji shows which share of one dollar of the domestically produced commodity i industry j makes. Two identities link these matrices and vectors:

$$\boldsymbol{q} = \boldsymbol{B}\boldsymbol{g} + \boldsymbol{e}, \tag{18a}$$

$$\boldsymbol{g} = \boldsymbol{W}\boldsymbol{q}. \tag{18b}$$

The first identity says that the dollar amount of each domestically produced commodity equals the sum of the dollar amount of that commodity that the different domestic industries use as intermediate goods plus the dollar amount of final uses of that commodity. Note again that final uses are for domestic private and public consumption, domestic investment, and net exports. The second identity says the dollar output of each industry equals the sum of that industry's contribution to the outputs of the different domestically produced commodities.

To eliminate $\boldsymbol{q}$ from these identities, we substitute (18b) into (18a) to obtain $\boldsymbol{q} = \boldsymbol{B}\boldsymbol{W}\boldsymbol{q} + \boldsymbol{e}$. We then solve this for $\boldsymbol{q}$ and substitute the result back into (18b). This gives:

$$\boldsymbol{g} = \boldsymbol{W}(\boldsymbol{I} - \boldsymbol{B}\boldsymbol{W})^{-1}\boldsymbol{e}. \tag{19}$$

$\boldsymbol{R} \equiv \boldsymbol{W}(\boldsymbol{I} - \boldsymbol{B}\boldsymbol{W})^{-1}$ is called the industry–by–commodity total requirements matrix. Rows are associated with industries and columns with commodities. Entry ji shows the dollar value of industry j's production that is required, both directly and indirectly, to deliver one dollar of the domestically produced commodity i to final uses including net exports.

Let $\boldsymbol{v}$ denote the $(1 \times n)$ vector of industry value added per unit of industry output, which is easily calculated from the IO Tables by dividing industry value added by industry output. To obtain the value added, $\boldsymbol{va}$, that is generated by the domestically produced final expenditure vector, $\boldsymbol{e}$, we multiply $\boldsymbol{R}$ (as defined either in (17) or in (19)) with $\boldsymbol{e}$:

$$\boldsymbol{va} = <\boldsymbol{v}> \boldsymbol{Re},$$

where $<\boldsymbol{v}>$ denotes the diagonal matrix with vector $\boldsymbol{v}$ in its diagonal. It is important to realize that this formula works for any domestically produced final expenditure vector, and so in principle we could use it for the domestically produced final consumption vector. However, we don't know this vector because we do not know the share of imports that is consumed. Instead, we only know the final consumption vector, $\boldsymbol{e}_c$. Component i of this vector reports the final consumption of commodity i, which may either be produced domestically or be imported. Assuming that imported commodities are produced with the same input requirements as in the U.S., we can use the total requirement matrix together with the vector $\boldsymbol{e}_c$. This gives us the consumption value added vector we are looking for:

$$\boldsymbol{c} = <\boldsymbol{v}> \boldsymbol{Re}_c. \tag{20}$$

In words, the vector on the left–hand side reports the value added in the different industries that is generated by the final consumption expenditure vector $\boldsymbol{e}_c$. Aggregating the components of this vector into our three broad sectors agriculture, manufacturing, and services gives us the consumption value added used in the text.

Appendix D: Results Without Government Services

Here we offer the results of our estimation exercises when we exclude government services. Table 7 reports the results. For convenience, the first and the third column repeat the results with government services from Table 1 and Table 3.

Table 7: Results With and Without Government Services

	Final Consumption Expenditure with	Final Consumption Expenditure without	Consumption Value Added with	Consumption Value Added without
σ	0.94**	0.49**	0.06	0.04
	(0.06)	(0.05)	(0.09)	(0.11)
$\bar{c}_a$	−1175.73**	−638.58†	−121.92**	−97.03**
	(28.87)	(395.26)	(9.28)	(13.27)
$\bar{c}_s$	12867.89**	2726.37**	3474.89**	4777.55**
	(3508.55)	(395.26)	(693.05)	(813.90)
ω_a	0.02**	0.24**	0.001	0.002*
	(0.0002)	(0.02)	(0.001)	(0.001)
ω_m	0.16**	0.20**	0.157**	0.145**
	(0.01)	(0.005)	(0.038)	(0.007)
ω_s	0.82**	0.56**	0.842**	0.852**
	(0.02)	(0.01)	(0.039)	(0.008)
AIC	−895.44	−667.04	−835.35	−774.46
$RMSE_a$	0.004	0.199	0.005	0.006
$RMSE_m$	0.009	0.013	0.011	0.016
$RMSE_s$	0.010	0.020	0.011	0.008

Robust standard errors in parentheses

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$