Economies of Scale in the Household:
Evidence and Implications from the American Past

Trevon D. Logan
The Ohio State University and NBER

April 5, 2007

Abstract

Household economies of scale arise when households with multiple members share public goods, making larger households better off at lower per capita expenditures. Research into household scale economies has yet to consider how household economies of scale change over time. I use American household expenditure surveys, covering 1888 to 1935, to produce the first comparable historical estimates of household scale economies in consumption. Scale economies in clothing, entertainment, and housing declined from 1888 to 1935, consistent with market expansion and increasing substitutes for these expenditures over time. Households in the past had fewer scale economies in food than today, however, exactly the opposite of what theory would predict and deepening a well known puzzle in the literature. Overall, I find that scale economies changed significantly from 1888 to 1935 for all expenditure categories. I then consider the implications of changing scale economies for estimates of real income. Previous estimates of CPI bias based on Engel curves do not account for changing scale economies in the household, and this can lead to omitted variable bias. My estimates of the annual rate of CPI bias from 1888 to 1935 are reduced by at least 25% once changing scale economies are accounted for, consistent with household economies of scale having a large, material effect on estimates of real income.

JEL Classifications: D1, E3, J1, N3

[0]Mailing Address: Department of Economics, The Ohio State University, 410 Arps Hall, 1945 North High Street, Columbus, OH 43210. E-Mail: logan.155@osu.edu. I thank Louis Cain, Dora L. Costa, Matthew S. Lewis, Audrey L. Light, Muna S. Meky, Joel Moykr, Paul Rhode, Gianni Tonolio and seminar participants at Ohio State, Indiana, McGill, Northwestern, and Duke for helpful discussions. Yasin Akcelik and Yunhui Tan provided excellent research assistance. The usual disclaimer applies.

1 Introduction

In one example of economies of scale, households with multiple members are able to achieve the same standard of living at lower per capita expenditures on public goods than smaller households. For example, if two adults unite to form one household, the couple will be better off as they can share public goods— such as housing— lowering per capita expenditure on them. The couple can funnel these savings into increased consumption of public goods— more spacious housing than either could afford separately— making each member better off than they were when they lived by themselves. The basic idea has strong intuitive appeal; living standards for households of different sizes could be equated with lower per capita expenditures for larger households who are able to economize on public goods. Measuring the distribution of income, the extent of poverty and deriving poverty thresholds require an accounting for these household economies of scale in consumption. For these reasons, measurement of economies of scale is fundamental to the measurement of living standards.

One difficulty with household economies of scale in consumption is that it can be difficult to identify the size of the scale economy.[1] While economies of scale are a function of the size of the household, the demand for goods and services generally derives from both household size and composition. When researchers turned their attention to economies of scale in the household, they overcame the problem of detection and measurement by concentrating on adult-only households, as in Nelson (1988). Measuring economies of scale for households where all members have identical tastes and are treated equally, however, will not aid in the measurement of poverty and the distribution of income, where compositional heterogeneity matters.

A second difficulty lies in the income and substitution effects that scale economies have on private goods. While both the income and substitution effects imply increased consumption of public goods, the income and substitution effects work in opposite directions for private goods, and it is not possible to know which effect will dominate unless we know more about the private good. Private goods which have low own– and cross price elasticities and sizable income elasticities are goods whose consumption is likely to increase with household size. The increased consumption of those private goods could be

[1] This difficulty is related to the more general issue of the proper modeling of demographic variables in demand analysis, although the primary goal of that literature was the creation of equivalence scales that were consistent with utility theory (Gorman 1976, Deaton and Muellbauer 1980, Pollack and Wales 1981, Lewbel 1985, Blundell, et. al. 2003).

taken as a measure of household economies of scale. Using this insight, Deaton and Paxson (1998) suggest that food would be a good choice for a private good where the income effect will dominate because food has few substitutes, and they also argue that substitution away from food will be even less likely for households close to subsistence. This would imply that food expenditures should increase with household size– the savings realized on public goods would be funneled to food, a necessity which has few substitutes. Rather than increasing with household size (with per capita expenditure held constant), Deaton and Paxson find that expenditure per capita on food falls with household size using a number of contemporary household expenditure surveys. After considering and dismissing a number of possible explanations, Deaton and Paxson conclude that their empirical results are a puzzle.

A literature analyzing this "food puzzle" has developed (Perali 2001, Horowitz 2002, Gibson 2002, Gan and Vernon 2003, Deaton and Paxson 2003, Vernon 2005), but research on household scale economies in consumption has yet to look at how or if scale economies change over time. Just as knowledge about scale economies at a point in time tells us about poverty and inequality, knowledge about changes in scale economies is useful to analyze changes in living standards over time. Even more, knowledge about changes in scale economies is important since average household size has changed over time. Indeed, Pistaferri, et al. (2005) use the concept of substitution between public and private goods to explain declines in household size over time. If the effect of household size on demand changes along with average household size we would need to know the extent of both to accurately measure trade-offs between public and private goods in the household. Lastly, time series information on economies of scale could allow us to resolve puzzles in the literature by eliminating potential explanations for the "food puzzle" and providing evidence of trends in economies of scale more generally.

With these ideas in mind, this paper has three related goals: (1) To estimate household economies of scale in the American past to see how they have changed over time. This paper provides the first comparable historical estimates of household economies of scale in consumption. (2) To use estimates of economies of scale in the past to analyze the empirical "puzzles"– seeing which expenditure categories are consistent with theoretical predictions and which are not. (3) To consider the implications of changing economies of scale on measures of living standards in the past, particularly the measurement of real income.

Using historical household surveys from the United States, covering the period 1888 to 1935, I estimate household scale economies for food, clothing, entertainment and housing. Scale economies in food show that the "food puzzle" holds for the past as well. In fact, I find that households in the past had *fewer* scale economies in food than today, exactly the opposite of what theory predicts. This finding deepens the "food puzzle" and cast serious doubt on whether food should be considered a private good for the measurement of household economies of scale. Furthermore, I find that existing explanations for the "food puzzle" do not explain the historical results. The other expenditure categories are consistent with theoretical predictions. Scale economies in clothing declined over the period considered, and are the same in 1990 as they were in 1935. Scale economies in entertainment also fell from 1888 to 1935. Housing expenditures decreased significantly for larger households in 1888, but by 1935 household size had a substantially smaller effect on housing expenditures, which suggests decreasing public good properties for housing over time. In general, I find that scale economies changed significantly over time for every expenditure category considered.

I also analyze a key implication of changing scale economies in the household– the measurement of real income. Costa (2001) and Hamilton (2001) have used Engel curves to estimate Consumer Price Index (CPI) bias in the past. Their methodology, however, does not estimate CPI bias independent of changes in household scale economies. I show that the CPI bias methodology, which attributes differences in food shares over time not explained by household characteristics and relative price changes to CPI bias, implicitly assumes that scale economies are unchanged over time. I modify the CPI bias estimation methodology, using an Engel curve that allows the scale economy to vary over time independent of CPI bias. When I estimate CPI bias in a way that controls for changing scale economies the CPI bias estimates, while still statistically significant, are reduced by at least 25%. This result suggests that changes in the effect of household size on demand play a material role in the measurement of real income.

The paper unfolds as follows. The next section summarizes the theoretical model and highlights empirical predictions about scale economies in household consumption in the past and over time. The third section outlines the empirical methodology. The fourth section discusses the estimates of scale economies in the household and discusses the results in light of the "food puzzle" described above. The fifth section analyzes CPI bias, and presents estimates of CPI bias that control for changing scale economies over time. The final section concludes.

2 Measuring Household Economies of Scale

There are two popular methods in the literature for measuring economies of scale in the household. The oldest and best known is the Engel method, while recent literature uses the Barten model to measure economies of scale. The Engel method has been preferred since it is easy to compute, but, as will be shown below, the Barten model is the superior approach for generating theoretical predictions about economies of scale in the household, although this claim has been debated in light of the "food puzzle" (Perali 2001, Gan and Vernon 2003). Below, I review both methods, show why the Barten model will be employed here, and highlight the theoretical predictions about household economies of scale.

2.1 Engel Measures of Household Economies of Scale

In Engel measures of economies of scale the scale economy is simply the difference in per capita expenditures between two households (the per capita expenditure of the smaller household minus that of the larger household) who devote the same share of total expenditure to food. Following Engel's Second Law, two households with identical budget shares devoted to food are equally well off if the foodshare is an indicator of the standard of living. Taking w to be food's share of the budget, x as total expenditure, and n as household size, the Engel method uses $\left(\frac{x}{n}\right)_{\overline{w_k}} - \left(\frac{x}{n}\right)_{\overline{w_j}}$, where $\overline{w_i} = w\left(\frac{x}{n_i}, n_i\right)$, $j > k$ and $\overline{w_k} = \overline{w_j}$. Figure 1 gives a graphical description of the Engel method. This method almost always gives positive values for scale economies– larger households have lower per capita expenditures.

The Engel method implies, however, that the larger household has lower per capita *food* expenditures than the smaller household as well. At a constant budget share of food, lower per capita expenditures by the larger household imply lower per capita food expenditures for the larger household as well, $\left(\frac{x}{n}\right)\overline{w_k} > \left(\frac{x}{n}\right)\overline{w_j}$. This assumption disagrees, directly, with the notion that there are economies of scale in the household since consumption of the private good (food) is lower for larger households. The Engel method does not give us theoretical justification for its measure of economies of scale, or show us how this measure is related to scale economies in the household in a tractable way. As such, we cannot use the Engel method to estimate economies of scale in the household since it is not a theoretically grounded measure of scale economies.[2]

[2] Nicholson (1976) has shown that Engel's method cannot be used to calculate child equivalence scales, but Perali

2.2 The Barten Model

The Barten model, developed first by Barten (1964) and extended to the analysis of scale economies in the household by Muellbauer (1977), Pollak and Wales (1980, 1981), Deaton and Muellbauer (1986) and Nelson (1988), generates theoretical predictions about scale economies. As such, it has served as the model used by Deaton and Paxson (1998) and the basis for those who have attempted to resolve the puzzle they discovered and others more generally concerned with public and private goods in the household (Perali 2001, Horowitz 2002, Gan and Vernon 2003, Deaton and Paxson 2003, Pistaferri, et al. 2005, Vernon 2005). In the two good Barten model a household of size n allocates expenditure on two goods. For convenience, one good is completely private, f (food), and the other is completely public, h (housing). The household maximizes the utility function

$$\max_{q_f, q_h} n\upsilon\left(\frac{q_f}{\phi_f(n)}, \frac{q_h}{\phi_h(n)}\right)$$

where $\upsilon(.)$ is the utility function, q_f and q_h are the quantities of food and housing respectively, and $\phi_f(n)$ and $\phi_h(n)$ are scaling functions for the economies of scale realized for food and housing respectively. The scaling function converts the size of the household (n) to its effective unit for both housing and food. In per capita terms the household budget constraint is

$$p_f\left(\frac{q_f}{n}\right) + \left(\frac{p_h}{n}\right) q_h = \frac{x}{n}$$

Maximization of the utility function subject to the budget constraint gives a per capita demand function for food

$$\frac{q_f}{n} = \frac{\phi_f(n)}{n} g_f\left(\frac{x}{n}, \frac{p_f \phi_f(n)}{n}, \frac{p_h \phi_h(n)}{n}\right)$$

where $g_f(\cdot)$ is the per capita demand function for food, per capita expenditure on food is

$$\frac{p_f q_f}{n} = \frac{p_f \phi_f(n)}{n} g_f\left(\frac{x}{n}, \frac{p_f \phi_f(n)}{n}, \frac{p_h \phi_h(n)}{n}\right) \tag{1}$$

What we would like to know is how much per capita food expenditure changes for a given increase in household size. Taking logs of equation (1) above and taking the derivative with respect to household size yields

(2001) argues that the puzzle is a problem of functional form. He concedes, however, that more theoretical research on household economies of scale is needed because "the literature on household economies of scale is not fully developed in the sense that the concept of economies of scale in the household does not have a close analog to the traditional concept defined in production theory." (p. 19) This issue is discussed futher in section 4.6.

$$\frac{\partial \ln (p_f q_f/n)}{\partial \ln (n)} = \left[1 - \frac{\partial \ln \phi_h (n)}{\partial \ln (n)}\right] (\epsilon_{fx} + \epsilon_{ff}) - \left[1 - \frac{\partial \ln \phi_f (n)}{\partial \ln (n)}\right] (1 + \epsilon_{ff}) \quad (2)$$

where ϵ_{ff} is the own price elasticity of food and ϵ_{fx} is the income elasticity of food.[3] In order for per capita food expenditure to increase with household size it must hold that

$$\left[1 - \frac{\partial \ln \phi_h (n)}{\partial \ln (n)}\right] (\epsilon_{fx} + \epsilon_{ff}) > \left[1 - \frac{\partial \ln \phi_f (n)}{\partial \ln (n)}\right] (1 + \epsilon_{ff}) \quad (3)$$

When food has few substitutes, meaning that ϵ_{ff} is small in absolute value, and if food has fewer scale economies than housing, meaning that $\frac{\left[1 - \frac{\partial \ln \phi_f (n)}{\partial \ln (n)}\right]}{\left[1 - \frac{\partial \ln \phi_h (n)}{\partial \ln (n)}\right]}$ is small, then we would expect the derivative given in (2) to be greater than zero.[4] This is essentially asserting that the savings obtained by economizing on the public good are used to increase consumption of the private good, where the income effect dominates the substitution effect for the private good because there are few substitutes for it. In this way, the Barten model gives us theoretical grounds upon which we can hypothesize about when, where, and for what goods we would expect economies of scale to be present.

To summarize, the Barten model predicts that:

- Holding per capita expenditure constant, the share of the budget devoted to f (the private good) will increase with household size if f has few substitutes and a sizable income elasticity.
- If it is true that poorer households have fewer substitutes for expenditures on f, such that ϵ_{ff} is smaller for poorer households, the increase in expenditures on f will be greatest for the poor.
- Over time, as the market expands and more substitutes for expenditure on particular private goods are available, ϵ_{ff} will increase and economies of scale in private goods should decrease.

The Barten model also has implications for public goods and their scale economies over time:

- Holding per capita expenditure constant, the share of the budget devoted to h (the public good)

[3] Note that if a good is purely public that $\left[1 - \frac{\partial \ln \phi_i (n)}{\partial \ln (n)}\right] = 1$ because there would be no change in the scaling function for an increase in household size. If a good is purely private then $\left[1 - \frac{\partial \ln \phi_i (n)}{\partial \ln (n)}\right] = 0$ because the scaling parameter would change exactly as much as household size.

[4] Another way to rewite the inequality is $\frac{(\epsilon_{fx} + \epsilon_{ff})}{(1+\epsilon_{ff})} > \frac{\left[1 - \frac{\partial \ln \phi_f (n)}{\partial \ln (n)}\right]}{\left[1 - \frac{\partial \ln \phi_h (n)}{\partial \ln (n)}\right]}$. When the income elasticity of the private good is high (approaching unity) a private good that is only slightly more private than the public good will suffice for the condition to hold.

will decrease with household size. This would be the source of cost savings funneled towards private goods.

2.3 Household Economies of Scale in the Past

We can use our knowledge of the past to generate further predictions about economies of scale and how we believe they would change over time. While we know that the predictions of the Barten model have been rejected for food in contemporary populations, there are several reasons to believe that the model should hold for food and other private goods in the past. Historically there were fewer substitutes for food, and it was not possible to substitute food preparation for expenditure to the extent that it is today.[5] As such, it is reasonable to expect ϵ_{ff} to be particularly small in the past. Similarly, the demand for food in the past was greater than it is today, and there is evidence that demand for nutrition in the past was greater than it is in developing nations today.[6] Logan (2006) estimates food income elasticities above 0.8 for the United States in the late nineteenth century. If the income elasticity of food, ϵ_{fx}, was very large in the past it would be even more likely food consumption would increase with household size. All of this implies that we should expect for food to behave in a manner consistent with the Barten model in the past– holding per capita expenditure constant food expenditure should increase with household size.

Furthermore, food is not the only private good that can be tested against the predictions of the Barten model, and Horowitz (2002) has argued that food may not be the appropriate private good on theoretical grounds. It is therefore useful to estimate economies of scale for other household expenditure categories. Clothing and entertainment expenditures have each been considered private goods in the literature, although the degree to which each is private is subject to debate. Similarly, the Barten model has implications for public goods, and this can be tested by looking at housing expenditures to see if they are consistent with the predictions of the Barten model. In short, we can estimate economies of scale in the past for a number of different goods and see if the time trends would be consistent with the predictions of the Barten model for both public and private goods.

In this paper I estimate economies of scale with American household survey data covering 1888 to 1935. The survey data used here comes from three national consumer expenditure surveys taken in

[5] See Byington (1910), Cowan (1983) and Moykr (2000) for more on historical substitution between food expenditures and time preparation and Aguiar and Hurst (2005, 2006) and Vernon (2005) for more on contemporary substitution.

[6] See Logan (2005) for more on the comparison of calorie demand elasticities over time.

1888-1890, 1917-1919, and 1935-1936.[7] For each survey, I estimate the scale economies of consumption for food, clothing, entertainment, and housing. Each survey is a large national survey of consumer expenditures and these surveys are comparable insofar as they each detail household expenditures, income, and household composition. Similarly, each survey used a similar methodology, interviewing subjects in their homes, verifying expenditures where possible, and using consistent categories for products and services.[8] In addition, each survey has comprehensive demographic information on all household members, which allows us to measure the effect of household size separate from the effects of household composition. Because these surveys are broadly consistent over time, it is possible to derive time trends in household scale economies from them.

3 Empirical Strategy

The Barten model gives us conditions under which we would expect per capita expenditures on private goods to increase with household size. Knowledge of the demand for food in the past gives us further grounds to argue that per capita expenditure on food should increase with household size. Our empirical task is to determine if food expenditure per capita increased with household size, holding per capita income constant. The hypothesis is that

$$E\left(\frac{p_f q_f}{n_j} \mid j, \frac{x}{n}\right) > E\left(\frac{p_f q_f}{n_k} \mid k, \frac{x}{n}\right)$$

where $j > k$. Since we are conditioning on per capita expenditure, we can use the share of the budget devoted to food, w_f, since $w_f \equiv \frac{p_f q_f}{x} = \frac{\frac{p_f q_f}{n}}{\frac{x}{n}}$. In other words, at the same level of per capita expenditures, households that devote a larger share of their budget to food have larger per capita food expenditures by definition. This simplifies the task to estimating the Engel curve for each type of expenditure and testing whether the budgetshare increases with household size. In each survey, I take total annual expenditures on food, clothing, entertainment and housing and divide each separately by total annual household expenditure as the dependent variables (the budget shares) in the analysis that follows. Following the previous literature, I estimate economies of scale in four ways

[7] The surveys are the Department of Labor's Cost of Living of Industrial Workers in the United States and Europe (1888-1890), the Bureau of Labor Statistics' Cost of Living in the United States (1917-1919), and the Department of Labor's and Department of Agriculture's Study of Consumer Purchases in the United States (1935-1936).

[8] See the data appendix for more information on the data sources, including survey construction and summary statistics.

for robustness. Three of the methods allow for increasing flexibility of the Engel curve, and the fourth method addresses the problems of the endogeneity of the budget share with per capita expenditure.[9] I detail each estimation procedure below.

1. *Linear (Ordinary Least Squares)-* The first method estimates economies of scale with a linear Rothbarth Engel curve using ordinary least squares (OLS). This specification attempts to separate the effect of household size from household composition on the budget share. The regression takes the form

$$w = \alpha + \beta \ln\left(\tfrac{x}{n}\right) + \gamma \ln n + \sum_{k=1}^{K-1} \delta_k \left(\tfrac{n_k}{n}\right) + \zeta z + \varepsilon$$

   where w is the budget share, x is total expenditure, n is household size, k is a grouping of the household by age and sex (such that n_k/n is the fraction of the household belonging to demographic group k), and z is a vector of control variables including the fraction of the household that is employed, geographic (state) controls, and the industry that employs the head of the household.[10] The composition of the household is broken into 5-year age-sex categories up to the age of 25. The measure of economies of scale is γ, the effect of household size on demand.[11]

2. *Fourier Engel Curve-* The second method uses a Fourier functional form of the Rothbarth Engel curve, giving it greater curvature and flexibility. This regression takes the form

$$w = \alpha + \beta \ln\left(\tfrac{x}{n}\right) + \phi \ln\left(\tfrac{x}{n}\right)^2 + \sum_{j=1}^{3} \left[\vartheta_j \sin\left[j \ln\left(\tfrac{x}{n}\right)\right] + \xi_j \cos\left[j \ln\left(\tfrac{x}{n}\right)\right]\right] + \gamma \ln n + \sum_{k=1}^{K-1} \delta_k \left(\tfrac{n_k}{n}\right) + \zeta z + \varepsilon$$

   where once again the measure of economies of scale is γ.[12]

3. *First Differencing Method-* The third method allows the Engel curve to take the form of any continuous function, giving it the greatest flexibility. Following the method proposed by Estes and Honore (1995) and Yatchew (1997) we can difference out the Engel function if it is continuous and still obtain and unbiased (but not efficient) estimate of economies of scale. Note that if the

[9] Congruent with the previous literature, I assume that household size has a level effect on demand. Nonparametric tests of this assumption with historical survey data (as in Logan 2006) have shown that it is valid. See Deaton and Paxson (1998) for more on non-parametric analyis of household economies of scale, and Blundell, et. al. (2003) for more on non-parametric equivalence scales.

[10] As relative prices are fixed, the geographic controls would caputure the effect of differences in relative prices across space.

[11] Formally, estimates of the economy of scale would regress the log of the budgetshare on the log of per capita expenditure and household size, which would be consistent with the theory described earlier. Since the log of the budget share is simply a monotonic transformation of the budget share itself, all of the estimates here are qualitatively similar.

[12] In this analysis the range of the variable in the Fourier analysis must be less than 2π, or it must be rescaled. All of the per capita expenditure ranges considered in this paper have a range of less than 2π, and therefore do not need to be rescaled.

Engel function is continuous and the sample sufficiently large, sorting by per capita expenditure implies that $\Delta f\left(\frac{x}{n}\right) \to 0$, such that it can be omitted from the regression. The differencing is achieved by sorting the data by per capita expenditure and taking first differences such that $\Delta x = \left(x_{(i)} - x_{(i-1)}\right)$. The regression is

$$\Delta w = \alpha + \gamma \Delta \ln n + \sum_{k=1}^{K-1} \delta_k \Delta\left(\frac{n_k}{n}\right) + \zeta \Delta z + \nu$$

The measure of economies of scale is γ.

4. *Instrumental Variables-* The fourth method addresses the fact that budget shares and per capita expenditures are constructed from the same information and the errors of both may be correlated as well, which would lead to biased estimates of β. Also, since per capita expenditure and household size are also correlated, such errors would also lead to biased estimates of γ, the coefficient of interest. Even more, we do not know, a priori, which direction the bias would be in. Income, which is highly correlated with expenditures but measured independently of it, is a good candidate as an instrument for per capita expenditure. The instrumental (IV) estimates use the log of per capita income as an instrument for expenditure.

4 Estimates of Household Economies of Scale

In this section, I present estimates of household scale economies for food, clothing, entertainment and housing from 1888 to 1935. Testing another prediction of the Barten model, that poorer households should increase their consumption of private goods more than wealthier households, I estimate economies of scale by income quartile over time. I then consider possible explanations of the "food puzzle," and show that the results for food are inconsistent with theory. I conclude that the puzzle hinges on what it means to equate the welfare of households of different sizes and whether food expenditure is the measure of welfare that Engel intended.

4.1 Scale Economies in Food

Table 1 shows estimates of the scale economy of food estimated in 1888, 1917, and 1935. The coefficients in the table show the effect of household size on the budget share devoted to food (γ) for each estimation method. For example, linear (OLS) estimates show that if household size were doubled in

1888, the share of the food budget share would decrease by roughly 2.3%.[13] There are several items of interest in Table 1. First, the four estimation methods yield remarkably similar estimates of the scale economy with respect to food. Second, all of the estimates of the food scale economy are negative. Regardless of the year, the food share never increases with household size. This mirrors the result found in Deaton and Paxson, but the results in Table 1 deepen the puzzle to the extent that, in the past, the conditions under which food expenditures should increase with household size were stronger in the past. There were fewer substitutes for food expenditures than there are today and the income elasticity of food was high in the past. The failure of household size to be positively correlated with food expenditures in the past is truly a puzzle.

Even more troubling, there is no clear time trend from the estimates of the scale economy with food, a further contradiction of the Barten model. Although it is true that the estimates are statistically different from one another, there is no discernible time trend. The scale economy is least private in 1917, and closest to zero in 1888. The increase in the size and scope of the market should lead to greater substitutes for food, and the increasing number of substitutes should lead the scale economy to decrease over time, *ceteris paribus*. Although the scale economy does decrease from 1888 to 1917, it increases from 1917 to 1935. It is unclear what movements could be behind such a pattern.

As a further check of robustness, we can look at food consumption at and away from home. This information is not available in the 1888 survey, but it is detailed in the 1917 and 1935 surveys. Although meals out of the home were infrequent in the past, it is still possible that the negative coefficient on household size seen in Table 1 could reflect the fact that larger families consume fewer meals out of the home. If that was true, larger households would have lower food expenditures since meals out of the home are usually more expensive than those consumed in the home. Indeed, Aguiar and Hurst (2005, 2006) suggest that expenditures on food for households with large time endowments should be lower than households with smaller time endowments. Since meals outside of the home are costly, and do not take advantage of this economy of scale in household production, household with large time endowments should have lower expenditures on meals outside of the home, holding per capita expenditure constant. Economies of scale in food preparation would induce larger households to substitute towards home-produced meals in the past, and Vernon (2005) has suggested that this should hold for the present. Table 2 shows the coefficients on the log of family size for regressions

[13] More precisely, the effect of doubling household size would be $\gamma * \ln(2)$ (where $\ln(2) = .693$).

on food at and away from home. Food expenditures away from home increase with family size, and over time the effect grows larger. Food expenditures at home decrease with household size, and the trend from 1917 to 1935 is the same as it is for food expenditures in general. Meals away from home cannot explain the results in Table 1.[14]

The ultimate conclusions drawn from the scale economy of food are that it does not behave generally in the way that a private good consistent with the Barten model would and it does not have a clear time trend consistent with the Barten model. While we will consider the implications of this later, we can also compare these historical estimates of the scale economy to contemporary estimates. Deaton and Paxson's estimate for 1990 from the Consumer Expenditure Survey, -.008, is significantly lower than any of the historical estimates for the economies of scale in food consumption in Table 1. This further deepens the puzzle insofar as the scale economy in the past should be positive, yet empirically the historical estimates are all more *negative* than the contemporary estimates, where the conditions of the Barten model are *more* likely to hold.

4.2 Scale Economies in Clothing

A priori, clothing is a private good, and we would expect clothing expenditures to increase with household size. Table 3 shows estimates of the scale economy with respect to clothing expenditures. As with the estimates of the food scale economy, the estimation techniques yield similar estimates of the scale economy. Unlike the food scale economy, however, the clothing scale economy is always positive, consistent with the prediction for private goods in the Barten model. Increases in household size lead to increased clothing expenditure per capita.

The scale economy of clothing, like the scale economy of food, changes over time. Furthermore, the trend in the clothing scale economies is consistent with the predictions of the Barten model. As the market develops and there are more substitutes in the market for expenditures on the private good, the size of the scale economy should diminish over time, and this is exactly what happens with the scale economy of clothing.[15] The scale economy is 1888 is approximately 70% greater than the

[14] The results in Table 2 also contradict Gan and Vernon's (2003) finding that IV estimates for household size's effect on food expenditures out of the home are negative– each method here yields positive estimates.

[15] This could also be due to the real price of clothing declining over time, such that households devote expenditure on other private goods, or goods being placed on the market that substitute for home production, although these are not the only possibilities. Expansion of the market is difficult to quantify, especially during this time period. Adding to these complications is that fact that the consumer durables revolution also saw an expansion of advertising and the use of consumer credit– both of which can expand the market for goods by providing consumers with more information and by lowering the cost of ownership. For more on the economic history of this time see Olney (1991).

scale economy measured in 1917, and the scale economy in 1917 is approximately 40% larger than the scale economy measured in 1935. These differences are statistically significant. The conclusion from the clothing scale economy is that clothing appears to be a private good that behaves in a way that is consistent with the Barten model.

The clothing scale economy is also consistent with the Barten model over the twentieth century. The scale economy in clothing in 1990 was .0194. The estimate for 1935 is close to that range, but slightly higher. If the scale economy declined consistently over time, the effect of household size would have decreased by .00084 per year from 1888 to 1935, while it would have declined .00018 per year from 1935 to 1990. This suggest that scale economies in clothing decreased at a decreasing rate from 1888 to 1990. Such a trend is entirely consistent with the notion that the most likely substitutes for clothing appeared on the market between 1888 and 1935.

4.3 Scale Economies in Entertainment

Expenditures for entertainment may be private, if they are enjoyed by a certain segment of the household. The estimates of the scale economy in entertainment are consistent with entertainment behaving as if it were a private good. Table 4 shows the results. The economy of scale estimates for entertainment are positive for every year considered. As with food and clothing, the estimates for entertainment's scale economy change from year to year, and these differences are statistically significant. From the OLS results, the scale economy estimate in 1888 is 75% larger than the 1917 estimate, and the 1935 estimate is approximately 25% larger than the 1917 estimate.

As with food, the entertainment scale economy does not follow any particular time trend. The estimates in 1888 are the largest, and the 1935 estimates are larger than the 1917 estimates. So although entertainment appears to be a private good in the Barten model sense, it does not behave with any clear time trend as predicted by the Barten model. Overall, however, the changes in the entertainment economies of scale, while statistically significant, are not very different qualitatively–doubling household size would not increase entertainment's share of the budget by 1%, and usually less than half of that. This may be due to the fact that what comprises entertainment changes from survey to survey, and of all the expenditure categories considered it is the most difficult one to construct a time consistent estimate for.[16] Entertainment may incorporate some of the items that

[16] See the appendix for more on the expenditures that comprise entertainment in each survey.

would be substitutes for entertainment in a previous or subsequent survey. Due to the difficulty of constructing a time-invariant measure of entertainment, and the expansion of entertainment and leisure options on the market during this time period, it is not clear how to interpret the lack of a trend in the entertainment scale economy.[17]

The size of the scale economy in entertainment in 1990 was .0087, which is larger than nearly any of the historical estimates, and suggests that entertainment has become an increasingly private good over time. This would be consistent with the growth of the entertainment industry in general, and the increasing segmentation of the entertainment industry over time. Part of the decrease in the scale economy in entertainment from 1888 to 1917 could be due to the rapid increase of entertainment products on the market during that time, particularly movies, radio, and recorded music, many of which were marketed to the family as a whole (Donohue 2003). The increasing scale economies since 1917 would be consistent with specialized, and therefore more likely private, entertainment options from the 1920s to today.

4.4 Scale Economies in Housing

Housing is a public good. As such, the estimates of scale economies for housing should be negative if housing is consistent with the Barten model. This is a check to see if the Barten model's predictions are consistent with a good on the other side of the public/private divide. Table 5 shows the scale economies in housing to be consistent with the Barten model. In every year, increases in household size are correlated with decreases in the share of expenditure devoted to housing. As with the other expenditure categories considered, the scale economy in housing does change over time. As Table 5 shows, the scale economy estimate in 1888 is substantially larger than the 1917 estimate (in absolute value), and the 1917 estimate is more than twice the size of the 1935 estimate (in absolute value). In fact, the changes over time in scale economies are most dramatic for housing. As with the other expenditure items, these differences are statistically significant.

Even more, housing behaves in a way that is consistent with the Barten model over time. Market expansion would lead to substitutes for housing expenditures in the market, such that the scale economy for housing would decline in absolute value over time as the income and substitution effects of savings on public goods diminish in size. The results of Table 5 are consistent with such a conjec-

[17] For more on the difficulties of capturing a time consistent measure of entertainment see Costa (1999).

ture, and they also imply, given the dramatic declines in the scale economy, that there were numerous substitutes for housing expenditures by 1935. Improvements in the availability and performance of functions inside of the household such as heating, cooling, furniture, bedding, and appliances would have a strong impact on the presence of household scale economies for housing.[18]

Estimates for the scale economy in housing for 1990 was -.0532, which is larger than the 1935 value of the scale economy. A natural explanation for this reversal would be the increasing ease of home ownership, particularly with the advent of government insured mortgages in the United States. As is well known, these sorts of public policies led to a general expansion of the housing market.[19] All of this serves to decrease the cost of home ownership, and the income effect for this public good would be large as a result. This decrease in the cost of home ownership most likely outweighs the increasing number of substitutes to housing in the market, explaining the reversal of the trend.

4.5 Scale Economies and the Income Distribution

There is an additional robustness check that can be performed on the scale economies considered above, and acts as an additional test of the Barten model. The scale economy should also change as a function of income, regardless of the year in which it is measured. Even more, the Barten model predicts that if poorer households have fewer substitutes for the private good they will increase consumption of the private good more than wealthier households. To test this prediction, I estimated the size of the scale economy for each expenditure category by income quartile for each survey year.

Since food in general is not consistent with the Barten model, it is not clear how it should behave as a function of income. If the general property that poor households have fewer substitutes for food than wealthier households holds, then poorer households should have less negative estimates of scale economies in food than wealthier households. Table 6 shows this to be the case in every survey. Households in the first income quartile have household scale economies in food that are greater than the scale economies of households in the fourth income quartile. For clothing, the results are broadly consistent with the Barten model. In general, households in the first income quartile would increase their expenditures on clothing more than households in the fourth income quartile. With the general

[18] When furniture and appliances are added to the estimates of housing the results are less negative than for housing alone (see the appendix for the additional results), which is consistent with these other goods acting as substitutes for housing expenditure.

[19] One could also argue that the cost of housing, relative to all other goods, decreased over time. Such a calculation, however, would have to take into account the cost of housing versus the cost of homeownership, where credit markets decrease the cost of homeownership but not the cost of housing itself.

pattern of declining scale economies in clothing over time, however, the differences narrow, and by 1935 there are not statistically significant differences in the size of the clothing scale economy by income quartile.

In contrast to these findings, poorer households have the lowest increase in expenditures for entertainment with household size. This would be entirely consistent with entertainment being a luxury that poor households can ill afford, or poorer households choosing more public versions of entertainment. Housing expenditures, however, are more difficult to rationalize. While the 1888 results are consistent with poorer households having the largest scale economies in housing, the 1917 and 1935 results suggest that wealthy households have the largest scale economies in housing. This could be due to poorer households having more cheap substitutes to housing expenditures in the past such as bedding, appliances, and the like.[20] The results do show, however, that for the same income quartile, the scale economy in housing decreases over time. Overall, the income distribution results are broadly consistent with the predictions of the Barten model, such that the cross-section predictions about scale economies and the income distribution are supported by the data in most cases. Even conditioning on quantile of the income distribution, there is a large change in the scale economies over time.

4.6 Does the Past Help us Resolve the "Food Puzzle"?

Since the Barten model fails to hold in the past, when food expenditure was even more likely to increase with household size, we are left with two options: we can either abandon the idea that food is a private good, or we can try to uncover the reasons why food in particular does not fit to the predictions of the Barten model.[21] Specifically, how do we reconcile these historical results with the Barten model?

To begin, consider the classic Engel function $w\left(\frac{x}{n}, n\right)$ where w is the share of the budget devoted to food, x is expenditure, and n is the size of the household. We know from Engel's first law that $\frac{\partial w}{\partial \frac{x}{n}} < 0$ and empirically the results of Table 1 confirm that $\frac{\partial w}{\partial n} < 0$ both in the past and present. These two derivatives tell us that the share of the budget devoted to food is decreasing both in per capita expenditure and in household size. Figure 1 displays this. Engel measures of economies of scale, which were rejected earlier because of their lack of theoretical justification, use these two facts

[20] Just as the effect of household size on demand was less negative when furniture and other presumably public goods were included, the income dictribution results with this expanded definition of housing (not reported) do show some greater stratification by income quartile.

[21] It is important to note that these two options are not mutually exclusive.

to derive estimates of economies of scale. Indeed, if $\frac{\partial w}{\partial n} < 0$ and $\frac{\partial h}{\partial n} < 0$ (where h is the budget share devoted to housing), it must hold that $\frac{\partial[1-w-h]}{\partial n} > 0$ and more generally that $\frac{\partial[1-w]}{\partial n} > 0$ by Walras' law. This implies that food is indeed less private than everything else if the sign of the derivative is an indication of household economies of scale. While the finding that $\frac{\partial w}{\partial n} < 0$ may present a theoretical puzzle in light of the Barten model, this has been a feature of Engel curves since their inception, and one of the reasons that Engel argued that larger households could be welfare-equated to smaller households with lower per capita expenditures.[22] The true puzzle, it seems, is how such an assumption (or assertion) can be theoretically justified as being welfare equivalent.

This puzzle seems to turn on the idea of welfare itself. What does it mean, empirically, to welfare equate households of different sizes? If welfare equivalence is similar proportional expenditures, it would seem as if the Barten model fails and the Engel conjecture survives, although what this measure actually is remains unclear. If welfare equivalence is taken to mean equating consumption, however, the problem is more complicated—a model describing the household production of food is needed. It will be difficult to specify a model that would accurately describe the changes in household technology over this time period. While Engel assumed (implicitly) that households with more members would spend *less* per capita on food, perhaps because other factors in the food production process increase with household size, more than conjecture will be needed to generate the sorts of predictions that the Barten model gave for changes in economies of scale as a function of technological change.[23]

A factor that is obviously related to food production is time. Since household time is increasing in the size of the household, larger households can substitute time for expenditure on food while leaving food consumption unchanged. Aguiar and Hurst (2005, 2006) and Vernon (2005) have recently argued that the "food puzzle" could be explained by time as an input into the food production function. The problem for the present analysis is that while there are models of time use in a household production function, time is not the only other factor that would be important during this time period– electrification, the availability of natural gas, household refrigeration, transportation innovations, changes

[22] Consider a simple calorie production function $F(wx, O)$ where $F(*)$ is the calories available to the household and O factors other than food purchases involved in the production of the food that lands on the plates from which we eat (time, cooking technology, food transportation and storing technology, etc.). Two households of different sizes are equally well off if $\frac{F(wx,O)}{n_i} = \frac{F(wx,O)}{n_j}$ where $n_i \neq n_j$. If the food production function is strictly increasing in both arguments and the other factors of food production increases with household size, $\frac{\partial O}{\partial n} > 0$, then household expenditure on food must decrease with household size for households of different sizes to be equally well off– if not, larger households would be strictly better off.

[23] This has been argued in a similar context by Aguiar and Hurst (2005) and Gronau and Hamermesh (2001, 2006).

in agricultural technology, knowledge and dissemination of sanity food preparation techniques and other factors all play a role in the production of food at home from the middle of the nineteenth century onward. Even more, some factors would be complements to time input while others would be substitutes– not all technological change was labor saving for household production (Cowan (1983) and Moykr (2000)), and Pollack (1999) has argued that time-use data alone is insufficient to explain household technology. A question of whether the income or substitution effect dominates for food consumption will be embedded in a question of whether the complements or substitutes to time input win out in a food production model at a particular point in time. As such, we would have more unknowns than we could accurately estimate, and it will be impossible to identify the model.[24]

While time as an input in the food production function may explain parts of the puzzle for contemporary populations, the market for prepared foods was small in the past, such that the extent of substitutability would not be as great in the past as it is today. All households had to contribute significant time to food preparation in the past.[25] The results of Table 2 confirm that large households did not substitute towards home produced meals. Additionally, Cowan (1983) has shown that time input in household production most likely *increased* during the time period considered here, despite the advent of (presumably) labor saving technology in the household.[26] This could be due to redundant use of both new and old technologies in household production. Furthermore, while time is certainly an input into the production of food, time was also an input in the production of goods like clothing in the past.[27] Hours spent in household production would imply that "puzzles" would exist for other private goods in the past that have time as a significant input, or at least a case must be made for why food would be different from other time intensive processes.[28] In general, empirical

[24] While one could calibrate a model to describe this process, as Pistaferri, et. al. (2005) do, the plausibility of changing preferences, which would change the presumed subtitution between public and private goods, presents additional problems to the calibration. Imagine, for example, that the substitution between public and private goods is a function of income. Over time, changes in household economies of scale could reflect changes along the substitution function and/or shifts of the function itself. As such, models that hinge on the substitution between public and private goods will be unable to hold such parameters fixed. Even with such complications, Pistaferri, et. al. use a variant of the Barten model to show that growth in income does explain a non-trivial portion of the changing household economies of scale, consistent with the Barten model and with the results of Table 6.

[25] See Byington (1910), Oddy (1990), Kertzer and Barbagli (2002) and Logan (2005, 2006) for more on household production of food in the past.

[26] Greenwood, et. al. (2005) argue that technological change freed women from time spent on household chores, but Cowan (1983) notes that hours spent on household chores increased from 50 in the late nineteenth century to 56 in the early twentieth century, and only began to decline after WWII– well after the time period considered here. The relationship between changes in household production technologies and time use is still a subject of debate in the literature.

[27] Smith (1994), for example, shows that clothing expenditures for mothers were inversely related to fertility, but for men they were invariant.

[28] If this were true it would have to hold that poorer households would use time to substitute for expenditure more than

identification of the relevant factors from such models will be difficult given the historical data.

A further complication is that demand for certain types of household production may have changed significantly during this time period for reasons unrelated to changes in household technology. Moykr (2000) has argued that demand for cleanliness and sanitation of food increased as a result of knowledge of the germ theory of disease and the subsequent public health interventions and campaigns of the late nineteenth and early twentieth centuries. As such, demand for cleanliness and sanitation (including food preparation) changed, and this could explain the increased hours spent in household production. Similarly, technological changes like refrigeration would have changed food availability, which itself could change the time spent cooking and preparing food (although the direction is not clear). If demand and technology were changing simultaneously, it will be difficult to construct the proper counterfactual needed to estimate the labor savings that technological change brought to the household.

Another possibility is that the Barten model yields predictions which do not generalize to the more than two-good case. Both Horowitz (2002) and Gan and Vernon (2003) make points along this dimension, but Deaton and Paxson (2003) reformulate the Barten model to include multiple goods and still generate the same predictions regarding economies of scale in food.[29] Gibson (2002) has suggested that measurement error explains the "food puzzle" result, and he argues that the problem lay in the recall method use by expenditure surveys. Systematic measurement error in food, however, seems unlikely since all of the recall evidence (for food and other items) are measured at the same time. Given the fact that people make food purchases more often than other purchases, it seems unlikely that food expenditures would be measured with error and other expenditures would not be plagued by the same types of errors. It is not clear why food would not behave as a private good but the other expenditure categories would.

---

wealthier households– for example, the opportunity cost of time is lower for poorer households. (Also note, however, that consumption for households could not be equated without an inverse relationship between time input and income since expenditure increases with income.) As such, the results in Table 6 would have to show that expenditure on food and clothing would decrease with household size more for poorer households than wealthier households; Table 6 shows exactly the opposite.

[29] Deaton and Paxson (2003) give the generalization. Rather than $\left[1-\frac{\partial \ln \phi_h(n)}{\partial \ln(n)}\right](\epsilon_{fx}+\epsilon_{ff}) > \left[1-\frac{\partial \ln \phi_f(n)}{\partial \ln(n)}\right](1+\epsilon_{ff})$, the new formulation becomes $\widetilde{\sigma}(\epsilon_{fx}+\epsilon_{ff}) - \left[1-\frac{\partial \ln \phi_f(n)}{\partial \ln(n)}\right](1+\epsilon_{ff}) > \sum_{k\neq f}\widetilde{\epsilon}_{fk}\left(\left[1-\frac{\partial \ln \phi_k(n)}{\partial \ln(n)}\right]-\widetilde{\sigma}\right)$ where $\widetilde{\sigma}$ is the weighted average of the economy of scale factors for all non-foods and $\widetilde{\epsilon}_{fk}$ is the compensated elasticity of food demand with respect to the price of nonfood k. The summation measures the net increase or decrease in food demand caused by substitutions among the nonfoods when household size increases.

This look at historical evidence has sharpened the focus of how the puzzle may be resolved theoretically by considering household production models for food. Testing such models, however, will be difficult as both technology and preferences were likely changing over this time period. It is also unclear why a model is necessary for food, but the Barten model suffices to explain clothing, entertainment, and housing expenditures. All told, the "food puzzle" remains.

5 Changing Household Scale Economies and Real Income

The previous section established that household scale economies changed over time in the American past. Below, I highlight one important implication of these changing scale economies– estimates of real income in the past. I consider the implications of changing household scale economies to CPI bias for two reasons. First, this application fits with the argument made earlier that household economies of scale have important implications for the measurement of well-being over time. CPI bias would imply that income growth has been under or overstated, and the role that changing household scale economies would have in estimating CPI bias is an important and unexplored issue in the literature. Secondly, the implications of household scale economies to CPI bias is straightforward. Scholars have recently developed methods of estimating the extent of CPI bias with Engel curves, and since the estimates of scale economies are based on an Engel framework, the integration of the two follows with few caveats.

Below, I show how the methodology to estimate CPI bias with Engel curves assumes that household economies of scale are constant over time. This assumption, when considered against the empirical work of the previous section, could lead to biased estimates of CPI bias. I then estimate CPI bias with and without controls for changing household scale economies. The results show that CPI bias estimates, while still statistically significant, decline dramatically once changes in household scale economies are controlled for.

5.1 Measuring CPI Bias

There are several reasons to use consumer expenditure surveys to estimate CPI bias. As Hamilton (2001) argued, estimates of CPI bias for particular populations would tell us more about living standards than aggregate estimates of real income. Similarly, household surveys can act as a check against our estimates of real income by seeing in consumer expenditures agree with the estimates of price changes. Costa (2001) has argued that, due to data limitations, household surveys are likely the only source available to estimate CPI bias in the past. These arguments extend to other measures of living standards. For example, Deaton (2006) has recently argued that household surveys can perform the same function for estimates of Purchasing Power Parity (PPP). Using Engel curves to capture CPI bias hinges on the fact that Engel curves relate the budget share to real household income. If food's share of the budget moves more or less than we would predict given estimates of changes in prices (from, say, the CPI), then movements in real income, measured directly from the Engel curve, would tell us about mis-measurement in real income from sources such as the CPI.[30] Following Hamilton's (2001) and Costa's (2001) methodology for the measurement of CPI bias, we can capture CPI bias through Engel curves of the following form

$$w_{i,j,t} = \phi + \varphi \left[\ln\left(1+\Pi_{F,j,t}\right) - \ln\left(1+\Pi_{N,j,t}\right)\right] + \beta \left[\ln Y_{i,j,t} - \ln\left(1+\Pi_{j,t}\right)\right] + X'\theta + \sum_{t=1}^{T}\delta_t D_t + \sum_{j=1}\delta_j D_j + u_{i,j,t} \tag{4}$$

where w is the share of the budget devoted to food, Π are the cumulative price changes for food (F), non-food (N), and overall price changes, Y is household expenditure, X is a vector of household characteristics, and D is a set of dummies for year (t) and region (j).[31] Any CPI bias will be captured in δ_t, since two households with the same inflation adjusted expenditures and demographic composition should have the same shares of the budget devoted to food, since changes in relative prices are accounted for and income has been deflated. In this way, any changes in the Engel curve would be due to the mis-measurement of income over time– CPI bias. Implicit in this framework is the idea that Engel curves capture living standards and households of similar type should have the same welfare (the foodshare) over time, controlling for changes in prices and deflating income.

The issue of changing household scale economies concerns variables in the vector X, which does not contain any time varying covariates. A household of size n can be disaggregated into a finite number of distinct groups, such that $n = \sum_{i=1}^{N} n_i$. Both Hamilton and Costa estimate the regression in (4) with disaggregated household size

[30] Food is not the only expenditure category that can be used to estimate CPI bias. The methodology is entirely general, any expenditure category could be used. The availability of price indicies for food over long time spans leads food to be the primary expenditure category used since changes in relative prices must be controlled for.

[31] For the derivation of this equation see the appendix.

$$w_{i,j,t} = \phi + ... + \sum_{i=1}^{N} \theta_i n_i + ... + u_{i,j,t} \tag{5}$$

The functional form in (5) controls for household composition and the effects of household size simultaneously. The modeling implicitly assumes that not only composition, but also household size induce the same change on the food budget share over time. The Hamilton-Costa form cannot separate the effects of household size and composition. While one can change composition and not household size, one cannot change household size without changing the composition. It is impossible to estimate CPI bias while controlling for changes in the effect of household size on demand with the demographic modeling that Hamilton and Costa use. While one may want to assume that household composition has the same influence on demand over time, the empirical results above and the Barten model itself gives us strong evidence that the effect of household size on demand does vary over time. Since the effect of household size does vary with time, the estimates of δ_t that Hamilton and Costa attribute to CPI bias suffer from omitted variable bias because changing household scale economies are correlated with time.[32] Interacting all of the demographic variables with time would allow for both composition and size effects to change over time, but the stated goal of the Hamilton-Costa specification is to assume that composition effects are time invariant to highlight the role that changes in household scale economies have on estimates of CPI bias.[33]

Hamilton (2001) concedes that the method, which is indirect, attributes all movement in the Engel curve unexplained by the other covariates to CPI bias. If there are missing covariates from the regression, or if there are errors in variables, then estimates of CPI could be spurious, overstated, or understated depending on the particular specification problem. Fortunately, changing household scale economies can be incorporated into the Rothbarth model that was used to estimate the economies of scale. Since the Rothbarth regression is an Engel curve (and differs from the Hamilton Costa form only in being in per capita terms) it, too, can be used to estimate CPI bias while at the same time controlling for changing household scale economies.[34] Recall that in the Rothbarth model the

[32] Hamilton (2001) describes robustness checks to his specification, which included adding covariates, interacting income with other covariates, and an instrumental variables specification. None of these robustness checks interacted household demography measures with time.

[33] As Costa (2001) explains "If the O'Grady's in 1919 had the same total CPI-deflated houusehold expenditures as the Svensons in 1935 and both families had the same number of children, then I attribute differences in their food and recreation shares to CPI bias, controlling for changes in relative prices." (p. 1292)

[34] None of the assumptions needed to estimate CPI bias with Engel curves is violated by using the Rothbarth functional form. For the derivation and assumptions see the appendix.

regression took the form

$$w_f = \alpha + \beta \ln\left(\frac{x}{n}\right) + \gamma \ln n + \sum_{k=1}^{K-1} \delta_k \left(\frac{n_k}{n}\right) + \zeta z + \varepsilon$$

which disaggregates changes in household composition, $\left(\frac{n_k}{n}\right)$, and changes in household size, (n), on demand by design. The Rothbarth functional form can be easily augmented to estimate CPI bias by including terms for changes in relative price, deflating per capita income, and including variables for time and region.[35] Incorporating changes in the effects of household size on demand is straightforward. The regression now becomes

$$w_{i,j,t} = \alpha + \varphi\left[\ln\left(1 + \Pi_{F,j,t}\right) - \ln\left(1 + \Pi_{N,j,t}\right)\right] + \beta[\ln\left(\frac{x}{n}\right)_{i,j,t} - \ln\left(1 + \Pi_{j,t}\right)] + \gamma \ln n + \sum_{k=1}^{K-1} \delta_k \left(\frac{n_k}{n}\right) + \lambda\left[(\ln n) * D_t\right] + \sum \delta_t D_t + \sum \delta_j D_j + u_{i,j,t} \qquad (6)$$

so that changes in the Engel curve that derive from changing scale economies can be estimated separate from estimates of CPI bias. If changing scale economies have an effect over time on movements in the Engel curve (if λ is statistically and economically significant) then the specification in (6) will capture it, and estimates of δ_t will not suffer from the omission of changing scale economies in the household. This reformulation not only corrects for omitted variable bias due to changing household scale economies, but also formally test the proposition that scale economies have changed over time.

It is useful to distinguish the issue here, about estimates of CPI bias that may suffer from omitted variable bias, from the general issue of modeling demographic variables in demand analysis. I am concerned with the fact that changes in household size may have a separate, distinct effect on changes in demand, and should therefore be included in estimates of CPI bias that seek to control for changes in household composition. A separate issue is the modeling of demographic variables in demand equations in a way that is theoretically consistent. See Pollack and Wales (1981) and Lewbel (1985) for classic references on this issue, and Blundell, et. al. (2003) for the non-parametric case. While I make no claims about the proper modeling of demographic variables in demand systems, it is important to note that the discussion above is an important example of the modeling of demographic variables in demand systems more generally. If Engel based estimates of CPI bias are sensitive to the effects of household size on demand, estimates of CPI bias should purge the effects of household size from estimates of CPI bias.

[35] Note that regional (state) effects were controlled for in the estimates of scale economies discussed in the previous section.

5.2 CPI Bias and Household Scale Economies

To derive estimates of CPI bias, the surveys of two years must be combined so that changes in relative prices and income can be captured and controlled for. I estimate CPI bias from 1888 to 1917 and from 1917 to 1935 using the surveys described above. Table 7 shows estimates of CPI bias with and without controlling for changing household scale economies. As predicted estimates of CPI bias decrease once the changes in household size are included, and the differences in the estimates of CPI bias are striking, reduced by 25% or more. For 1888 to 1917, controlling for changes in household size reduces estimates of CPI bias by nearly three-fold. For 1917 to 1935, controlling for changes in scale economies reduced the estimate of CPI by nearly 25%.

Changing household economies of scale appear to have a large effect on estimates of real income. Figure 2 shows estimates of the CPI from 1888-1937 with and without household size bias corrections. These new estimates of CPI bias revise estimates of real consumer expenditures as well. Figure 3 shows CPI deflated real expenditures from 1900 to 1937, as well as household size corrected and uncorrected estimates. As the figure shows, the household size corrected estimates of real expenditures are closer to the CPI deflated real expenditure estimates than the uncorrected estimates. Controlling for changes in household size over time has a significant effect on estimates of CPI bias and real expenditure in the past. Estimates of CPI bias that do not control for changing household economies of scale overstate CPI bias.

Table 7 also shows two interesting facts. First, the inclusion of household size change over time does not impact the effect of income on demand, so household changes do not appear to have an effect that works directly through income itself. This is analogous to saying that changes in household size appear to be uncorrelated with deflated income. Secondly, changes in household size do not fully explain CPI bias. Even after controlling for changing household size, there is still a statistically significant effect of time that can be attributed to CPI bias. The inclusion of other time varying covariates (such as changing household compositional effects) further reduce estimates of CPI bias.[36] CPI bias estimates are sensitive to the modeling of household size and composition.

[36] When CPI bias estimates of the type described in this section were estimated with interactions on household composition (interacting the household shares with time), the effect of CPI bias decreased dramatically. For 1888-1917, the fully interacted model resulted in a "Year is 1917" coefficient estimate of 0.015, which was not statistically significant (p-value 0.11). For 1917-1935, the "Year is 1935" coefficient changes sign (0.029), but is still statistically significant (p-value 0.002).

6 Conclusion

This paper has documented changes in the effect of household size on demand from the late nineteenth to the middle of the twentieth century in the United States. Using historical household surveys and the Barten model to derive predictions about household economies of scale, I argued that the scale economies in the past would be stronger than they are today. The past, however, does not resolve the empirical puzzles with household scale economies. Congruent with previous results, food expenditure per capita decreased with household size in the past, although the conditions whereby food expenditure should increase with household size were strongest in the past. While the "food puzzle" remains, the time trend reveals that food has not behaved in any systematic fashion in the past, and that there were fewer scale economies in food in the past than at present. The results here, when combined with the cross-country evidence of Deaton and Paxson, cast serious doubt on the possibility that food was, or will be in the future, a private good consistent with the predictions of the Barten model. More troubling, the existing explanations for the "food puzzle" do not explain these historical results– as such, the "food puzzle" is only made more complicated by looking into the past.

Unlike food expenditures, clothing and entertainment behave in a manner more consistent with the Barten model for private goods, where clothing and entertainment expenditures increased with household size. The predictions of the Barten model for public goods holds for the past as well. Housing expenditures decreased with household size in the past. Similarly, the effect of household size on demand varied with household income at every point in time, although the slope of the relationship changed over time. In total, the effect of household size on demand for every consumption category changed significantly from the late ninteenth century to the middle of the twentieth century.

The changing effect of household size on demand has important welfare implications, and I explored here how estimates of CPI bias may suffer from omitted variable bias if changes in the effect of household size on demand are not accounted for. The results here suggest that estimates of CPI bias are reduced by at least 25% once changes in household size are controlled for. Much of what we may erroneously attribute to unaccounted changes in real income actually reflect the changing influence of household economies of scale on demand. As such, changes in household economies of scale have a large, material effect on welfare.

References

[1] Aguiar, Mark and Erik Hurst (2005). "Consumption versus Expenditure." *Journal of Political Economy* 113 (5): 919-948.

[2] Augiar, Mark and Erik Hurst (2006). "Lifecycle Prices and Production." Forthcoming, *American Economic Review.*

[3] Barten, Anton P. (1964). "Family Composition, Prices and Expenditure Patterns." in Peter E. Hart, Gordon Mills, and John Whitaker, eds. *Econometric Analysis for National Planning*. London: Butterworths.

[4] Blundell, Richard W., Martin Browning and Ian A. Crawford (2003). "Nonparametric Engel Curves and Revealed Preference." *Econometrica* 71 (1): 205-240.

[5] Byington, Margaret F. (1910). *Homestead: The Households of a Mill Town*. Philadelphia: W.M. Fell Co.

[6] Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, eds. (2006). *Historical Statistics of the United States*. New York: Cambridge University Press.

[7] Costa, Dora L. (1999). "American Living Standards: Evidence from Recreational Expenditures." NBER Working Paper No. 7148.

[8] Costa, Dora L. (2001). "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves." *Journal of Political Economy* 109 (6): 1288-1310.

[9] Cowan, Ruth S. (1983). *More Work for Mother*. New York: Basic Books.

[10] Deaton, Angus (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, MD: Johns Hopkins.

[11] Deaton, Angus (2006). "Purchasing Power Parity Exchange Rates for the Poor: Using Household Surveys to Construct PPPs." Unpublished manuscript, Princeton University.

[12] Deaton, Angus and John Muellbauer (1980). *Economics and Consumer Behavior*. New York: Cambridge.

[13] Deaton, Angus and Christina Paxson (1998). "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy* 106 (5): 897-930.

[14] Deaton, Angus and Christina Paxson (2003). "Engel's What? A Response to Gan and Vernon." *Journal of Political Economy* 111 (6): 1378-1381.

[15] Donohue, Kathleen G. (2003). *Freedom from Want*. Baltimore: Johns Hopkins.

[16] Estes, Eugena and Bo Honore (1995). "Partial Regression Using One Nearest Neighbor." Unpublished Manuscript, Princeton University.

[17] Gan, Li and Victoria Vernon (2003). "Testing the Barten Model of Economies of Scale in Household Consumption: Toward Resolving a Paradox of Deaton and Paxson." *Journal of Political Economy* 111 (6): 1361-1377.

[18] Gibson, John (2002). "Why Does the Engel Method Work? Food Demand, Economies of Size and Household Survey Methods." Unpublished Manuscript, University of Waikato.

[19] Gorman, William M. (1976). "Tricks with Utility Functions." in M.J. Artis and A.R. Nobay, eds., Essays in Economic Analysis: Proceedings of the 1975 AUTE Conference Sheffield. Cambridge: Cambridge University Press.

[20] Greenwood, Jeremy, Ananth Seshadri and Mehmet Yorukoglu (2005). "Engines of Liberation." *Review of Economic Studies* 72 (1): 109-133.

[21] Gronau, Ruben and Daniel Hamermesh (2001). "The Demand for Variety: A Household Production Perspective." NBER Working Paper No. 8509.

[22] Gronau, Ruben and Daniel Hamermesh (2006). "Time vs. Goods: The Value of Measuring Household Production Technologies." *Review of Income and Wealth* 52 (1): 1-16.

[23] Hamilton, Bruce W. (2001). "Using Engel's Law to Estimate CPI Bias." *American Economic Review* 91 (3): 619-630.

[24] Horowitz, Andrew W. (2002). "Household Size and the Demand for Food: A Puzzle Resolved?" Unpublished Manuscript, University of Arkansas.

[25] Kertzer, David, I. and Marzio Barbagli, eds. (2002). *The History of the European Family, Volume 2 Family Life in the Long Nineteenth Century.* New Haven, CT: Yale University Press.

[26] Lewbel, Arthur (1985). "A Unified Approach to Incorporating Demographic and Other Effects Into Demand Systems." *Review of Economic Studies* 52 (1): 1-18.

[27] Logan, Trevon D. (2005). "The Transformation of Hunger: The Demand for Calories Past and Present." NBER Working Paper No. 11754.

[28] Logan, Trevon D. (2006). "Nutrition and Well-Being in the Late Nineteenth Century." *Journal of Economic History* 66 (2): 313-341.

[29] Moykr, Joel (2000). "Why 'More Work for Mother?' Knowledge and Household Behavior, 1870-1945." *Journal of Economic History* 60 (1): 1-41.

[30] Muellbauer, John (1977). "Testing the Barten Model of Household Consumption Effects and the Cost of Children." *Economic Journal* 87: 460-487.

[31] Nelson, Julie A. (1988). "Household Economies of Scale in Consumption: Theory and Evidence." *Econometrica* 56 (6): 1301-1314.

[32] Nicholson, J. Leonard (1976). "Appraisal of Different Methods of Estimating Equivalence Scales and Their Results." *Review of Income and Wealth* 22 (1): 1-11.

[33] Oddy, D. J. (1990). "Food, Drink, and Nutrition." In F. M. L. Thompson, ed. *The Cambridge Social History of Britain 1750-1950, Vol. 2 People and their Environment.* New York: Cambridge University Press: 251-278.

[34] Officer, Lawrence H. (2006). "The Annual Consumer Price Index for the United States, 1774-2005" Measuringworth.com (data series at http://www.measuringworth.com/uscpi/).

[35] Olney, Martha L. (1991). *Buy Now, Pay Later: Advertising, Credit, and Consumer Durables in the 1920s.* Chapel Hill, NC: University of North Carolina Press.

[36] Perali, Federico (2001). "The Second Engel Law and Economies of Scale: An Empirical Puzzle to be Resolved." Unpublished Manuscript, University of Verona.

[37] Pistaferri, Luigi, Todd Schoellman and Michele Tertilt (2005). "Household Size and the Demand for Private Goods." Unpublished Manuscript, Stanford University.

[38] Pollak, Robert A. (1999). "Notes on Time Use." *Monthly Labor Review* 122 (August): 7-11.

[39] Pollak, Robert A. and Terrence J. Wales (1979). "Welfare Comparisons and Equivalent Scales." *American Economic Review* 69 (May): 216-221.

[40] Pollak, Robert A. and Terrence J. Wales (1980). "Comparison of the Quadratic Expenditure System and Translog Demand Systems with Alternative Specifications of Demographic Effects." *Econometrica* 48 (3): 595-612.

[41] Pollak, Robert A. and Terrence J. Wales (1981). "Demographic Variables in Demand Analysis." *Econometrica* 49 (6): 1533-1551.

[42] Smith, Daniel S. (1994). "A Higher Quality of Life for Whom? Mouths to Feed and Clothes to Wear in the Families of Late Nineteenth-Century American Workers." *Journal of Family History* 19: 1-33.

[43] Vernon, Victoria (2005). "Food Expenditure, Food Preparation Time, and Household Economies of Scale." Unpublished Manuscript, Fashion Institute of Technology, State University of New York.

[44] United States Bureau of Labor Statistics (1951). *Handbook of Labor Statistics: 1950 Edition.* Bulletin No. 1016. Washington, D.C.: Government Printing Office.

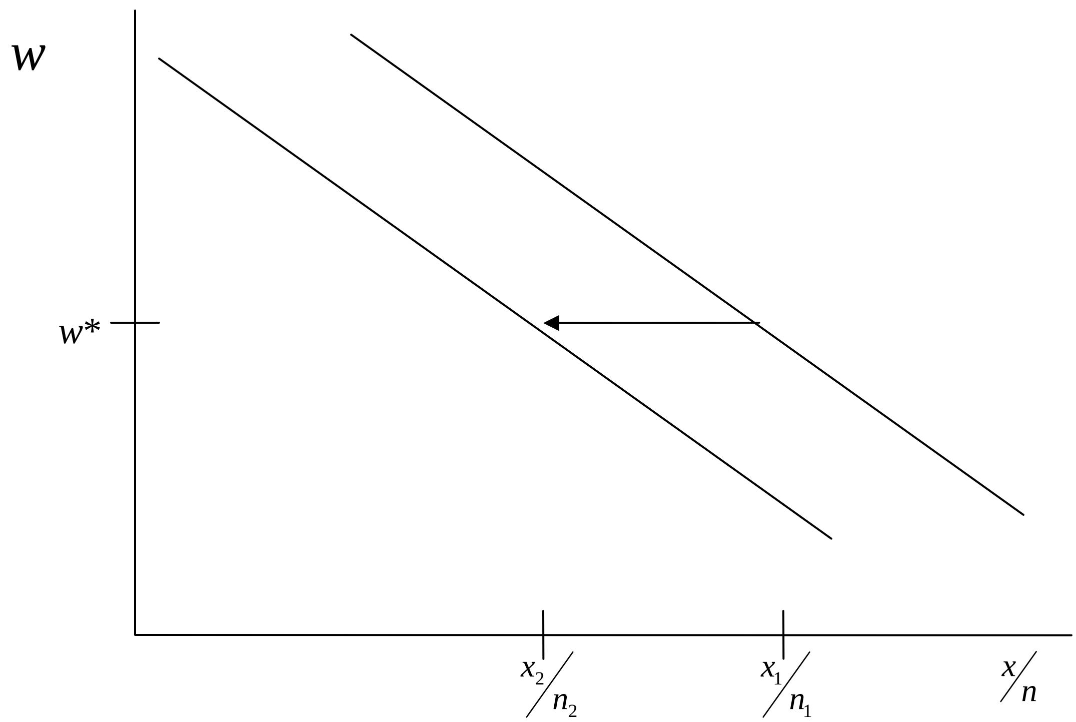
[45] United States Department of Labor (1986). *Cost of Living of Industrial Workers in the United States and Europe 1888-1890. ICPSR 7711.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

[46] United States Department of Labor (1986). *Cost of Living in the United States 1917-1919. ICPSR 8299.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

[47] United States Department of Labor and United States Department of Agriculture (1989). *Study of Consumer Purchases in the United States, 1935-1936. ICPSR 8908.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

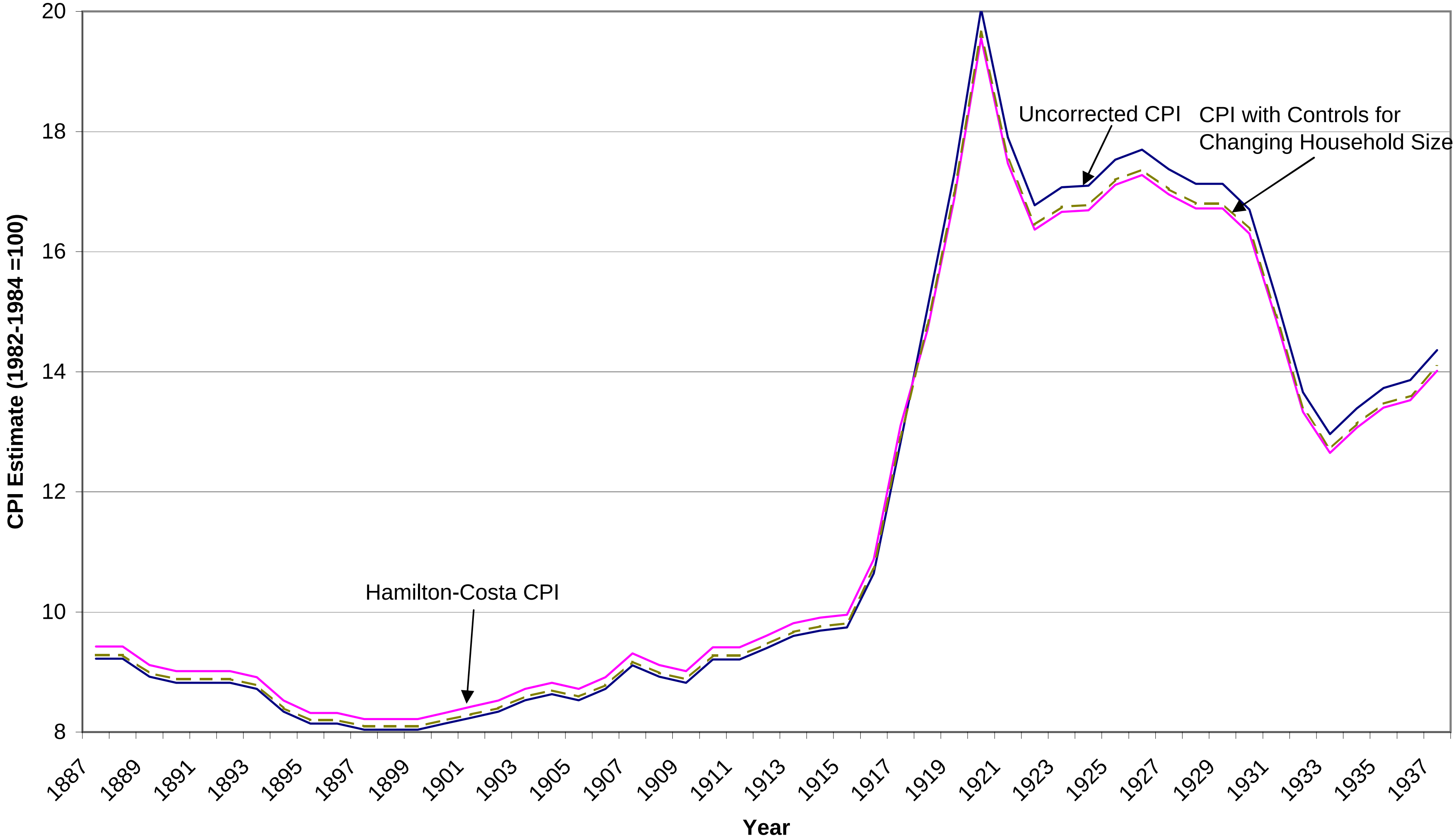
[48] Yatchew, Adonis (1997). "An Elementary Estimator of the Partial Linear Model." *Economics Letters* 57: 135-143.

Figure 1



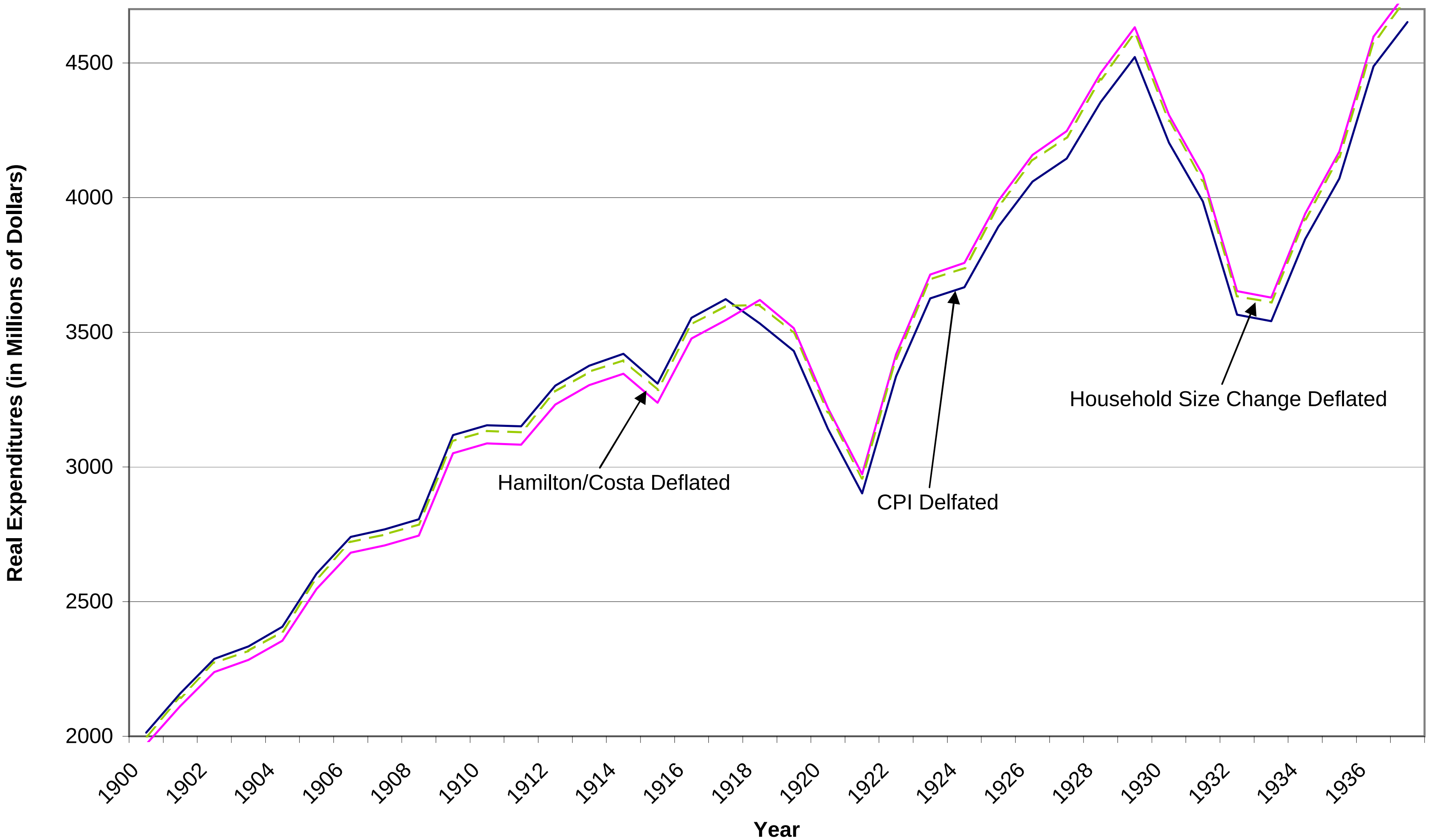
Engel Measures of Economies of Scale

Figure 2
Corrected CPI Estimates, 1887-1937, With and Without Controls for Changing Household Size



*Original CPI estimates come from Lawrence H. Officer (2006), "The Annual Consumer Price Index for the United States, 1774-2005" MeasuringWorth.com

Figure 3
CPI Deflated Real Expenditures, 1900-1937



*Real expenditure estimates come from Carter, et al. (2006) *Historical Statistics of the United States: Millennial Edition* .

Table 1
The Effect of Household Size on Food's Share of the Budget 1888-1935

	I	II	III
Method	1888	1917	1935
Linear (OLS)	-0.023	-0.091	-0.040
	(5.6)	(32.39)	(3.36)
Fourier	-0.022	-0.089	-0.051
	(5.16)	(31.67)	(4.32)
First Differencing	-0.026	-0.090	-0.054
	(6.27)	(31.97)	(4.27)
IV	-0.024	-0.083	-0.038
	(5.67)	(28.33)	(3.06)
N	6809	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household.
The dependent variable in each regression is the share of household expenditure devoted to food.
Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.

Table 2
The Effect of Household Size on Food's Budget Share by Consumption Type 1917-1935

	Food at Home		Food Away from Home	
	I	II	III	IV
Method	1917	1935	1917	1935
Linear (OLS)	-0.097	-0.046	0.006	0.008
	(34.35)	(4.5)	(5.84)	(2.39)
Fourier	-0.096	-0.044	0.007	0.008
	(33.72)	(4.32)	(6.08)	(2.27)
First Differencing	-0.096	-0.035	0.006	0.006
	(33.83)	(3.23)	(5.62)	(1.65)
IV	-0.089	-0.053	0.007	0.010
	(30.36)	(5.07)	(5.94)	(2.77)
N	12817	3534	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed,
the state of residence, the industry of the household head, and demographic shares of the household.
The dependent variable in each regression is the share of household expenditure devoted to food at home (I-II)
and away from home (III-IV).
Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.

Table 3
The Effect of Household Size on Clothing's Share of the Budget 1888-1935

	I	II	III
Method	1888	1917	1935
Linear (OLS)	0.068	0.040	0.029
	(21.78)	(17.83)	(5.23)
Fourier	0.069	0.044	0.028
	(21.78)	(19.31)	(4.96)
First Differencing	0.063	0.048	0.028
	(19.87)	(21.31)	(4.91)
IV	0.082	0.049	0.035
	(24.91)	(20.99)	(6.08)
N	6809	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household.
The dependent variable in each regression is the share of household expenditure devoted to clothing.
Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.

Table 4
The Effect of Household Size on Entertainment's Share of the Budget 1888-1935

	I	II	III
Method	1888	1917	1935
Linear (OLS)	0.007 (5.75)	0.004 (9.63)	0.005 (1.93)
Fourier	0.006 (5.1)	0.004 (10.0)	0.005 (1.94)
First Differencing	0.006 (4.82)	0.004 (10.5)	0.003 (0.89)
IV	0.010 (7.84)	0.005 (12.68)	0.008 (2.89)
N	6809	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household.
The dependent variable in each regression is the share of household expenditure devoted to leisure/entertainment/recreation.
Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.
See the data appendix for the expenditure categories that comprise entertainment.

Table 5

The Effect of Household Size on Housing's Share of the Budget 1888-1935

	I	II	III
Method	1888	1917	1935
Linear (OLS)	-1.538	-0.068	-0.029
	(3.54)	(22.41)	(3.63)
Fourier	-1.560	-0.067	-0.025
	(3.59)	(21.85)	(3.17)
First Differencing	-1.666	-0.065	-0.022
	(3.86)	(21.4)	(2.72)
IV	0.787	-0.069	-0.035
	(1.72)	(21.89)	(4.31)
N	6809	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household.
The dependent variable in each regression is the share of household expenditure devoted to housing.
Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.
See the data appendix for the expenditure categories that comprise housing expenditures.

Table 6
The Effect of Household Size on Budget Shares by Income Quartile, 1888-1935

Category	Income Quartile			
Food	1st	2nd	3rd	4th
1888	-0.007	-0.014	-0.032	-0.027
	(0.71)	(1.89)	(4.09)	(2.66)
1917	-0.075	-0.087	-0.105	-0.094
	(12.82)	(15.61)	(19.15)	(16.04)
1935	-0.026	-0.058	-0.072	-0.047
	(0.96)	(2.23)	(3.18)	(2.53)
Clothing	1st	2nd	3rd	4th
1888	0.075	0.065	0.057	0.054
	(10.43)	(11.34)	(9.51)	(7.16)
1917	0.050	0.053	0.039	0.022
	(12.0)	(12.33)	(8.8)	(4.0)
1935	0.022	0.033	0.032	0.036
	(1.97)	(2.93)	(2.82)	(3.28)
Entertainment	1st	2nd	3rd	4th
1888	-0.003	0.004	0.003	0.004
	(1.83)	(2.39)	(1.18)	(0.93)
1917	0.001	0.005	0.006	0.000
	(1.33)	(5.97)	(6.22)	(0.18)
1935	0.004	0.009	0.001	0.013
	(0.76)	(1.4)	(0.2)	(1.84)
Housing	1st	2nd	3rd	4th
1888	-5.259	-0.790	-0.457	0.062
	(4.26)	(0.95)	(0.6)	(0.08)
1917	-0.065	-0.060	-0.067	-0.082
	(11.4)	(10.25)	(11.02)	(11.53)
1935	-0.020	-0.034	-0.036	-0.053
	(0.86)	(1.66)	(1.89)	(2.97)

Note:
Each entry is the coefficient estimate for the log of household size for that income quartile, year pairing for that expenditure category, whose budget share is the dependent variable in the regression.
Each entry comes from a separate OLS regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household.
Robust t-statistics are listed under coefficient estimates in parenthses.

Table 7
Estimating CPI Bias with and without Controls for Changing Household Size, 1888-1935

	I 1888/1917	II 1888/1917	III 1917/1935	IV 1917/1935
Log Real Per Capita Expenditure	-0.132 (73.96)	-0.131 (73.11)	-0.159 (93.82)	-0.159 (93.89)
Log Household Size	-0.055 (23.44)	-0.069 (24.82)	-0.101 (35.20)	-0.099 (33.62)
Year is 1917	0.065 (34.83)	0.026 (5.70)		
Year is 1917 * Log Household Size		0.025 (9.36)		
Year is 1935			-0.091 (4.42)	-0.069 (3.17)
Year is 1935 * Log Household Size				-0.016 (3.15)
R-Squared	0.393	0.396	0.477	0.479
N	19626	19626	16467	16467
Cumulative CPI Bias	-0.633	-0.219	0.435	0.351
Annual CPI Bias	-0.022	-0.007	0.024	0.019
% Difference in annual bias without household control		288%		124%

Note:
The dependent variable in all regressions is the share of household expenditure devoted to food.
Robust t-statistics are listed under coefficient estimates in parentheses. The regressions above include relative price changes between food and non-food (by region), deflated household expenditure (be region), regional dummies, household demographics, and the fraction of the household employed.

See section 5 of the text for details on the functional form in the CPI bias regression.
See the appendix for the derivation of the CPI bias estimate.

Appendix

1 Further Evidence on Household Economies of Scale in the Past

For certain explanations of the "food puzzle" to hold the scale economies in food should display a specific time trend. If direct economies of scale in food cause the per unit price of food to be lower in larger households this effect should intensify over time as technology allows larger households to purchase and store large quantities of food, which are sold at cheaper prices in bulk. Since the effect of household size on demand is stronger in the past, however, such as explanation is unlikely to resolve the issue. Also, if income is distributed unequally within the household (where larger households have more unequal distributions than smaller households) food consumption could decline with household size and explain the "food puzzle."[37] If the Engel curve is relatively stable over time, however, the effect should attenuate and the scale economies in food should increase over time. While in general the effect does lessen over time, it does not do so in a consistent fashion (as it does for clothing), suggesting that the effect of household size does not act in a consistent way over time with respect to food. Overall, the evidence is not consistent with an inequality explanation.

There are additional tests of the "food puzzle" that other researchers have used to investigate (and potentially resolve) the puzzle. Gan and Vernon (2003) argued that the "food puzzle" resulted from improper specification of the predictions of the Barten model. In the two good model presented in this paper, food must be more private than housing for food expenditure per capita to increase with household size $\left(\text{where } \left[1 - \frac{\partial \ln \phi_h(n)}{\partial \ln(n)}\right] > \left[1 - \frac{\partial \ln \phi_f(n)}{\partial \ln(n)}\right]\right)$. Gan and Vernon argue that the generalization of this condition to a world with more than two goods would require food to be more private than all other expenditures, not only housing. They argue that all other goods, $(1 - w_f)$, will contain expenditure categories that are more public than food and others that are more private than food. Rather than testing to see if the foodshare increases with household size, they argue that food should be paired with other goods known to be more private than food so that the composite good (in their case food and clothing) will be more likely to be more private than all other household expenditures. They assert that such a test is closer to the predictions of the Barten model. In their paper they pair food and clothing expenditures and find that the food and clothing share is positively related to household size for their contemporary household surveys, and consistent with the idea that food and clothing are jointly private goods.[38] Unfortunately, the results do not hold for the historical household surveys used in this paper. Table A1 shows the results, where the food and clothing share is negative in 1917 and in 1935. This implies that the negative correlation of household size and the foodshare dominates the positive relationship between household size and the clothing share– such that food and clothing jointly behave more like public goods in the past.

Similarly, Gan and Vernon suggest that food should be tested directly against categories known to be more public than food, such as housing. They find that the share of food in food and shelter, $\left(\frac{f}{f+h}\right)$, is positively related to household size in their contemporary household surveys.[39] They take this as evidence that food is more private that housing, which they argue is in agreement with the predictions of the Barten model. Table A1 shows that this was not always the case– food in food and housing was negatively related to household size in the past. Given these results, Gan and Vernon's attempts to resolve the puzzle by grouping food with other items and comparing food to other goods known to be more public than food do not generate similar conclusions with historical evidence. Overall, the findings in regards to food help to eliminate some potential explanations, but do not resolve the food puzzle or the debates regarding the Barten model.

2 Estimating CPI Bias

Estimating CPI bias from Engel curves begins with a number of assumptions. Decomposing food and non-food expenditures into a price index and quantity index requires that food be additively separable in the household's utility function. Furthermore, there must be homotheticity in the subutilities of food and non-food. With

[37] This also requires that the Engel curve be concave.

[38] Their surveys come from the United States (1990 Consumer Expenditure Survey), South Africa (1993 Living Standards Survey), and Russia (1994-1998 Russian Longitudinal Monitoring Survey).

[39] As Gan and Vernon note, analyzing the share of food in food and housing requires that utility be separable with respect to food and housing and assumes that the household optimally allocates expenditures between food and housing and all other items.

these conditions the bias of non-food does not effect the foodshare through complementarities of substitutabilities through some unmodeled channel. Hamilton (2001) further notes that food is chosen because (1) it has an income elasticity that is different from unity, and therefore sensitive to the measurement of income, (2) it is non-durable and therefore not subject to stock and flow effects (this would, naturally, be stronger in the past than today), (3) food does not involve the troublesome "definitional problems" of other expenditure categories and (4) because the Working-Leser Almost Ideal Demand System (AIDS) has a functional form that has successfully estimated the demand for food. It is also important to note that the method requires the dependent variable to be the budget share for food– food consumption and expenditure are likely to contain CPI bias themselves.

Beginning with the Almost Ideal Demand System for food

$$w_{i,j,t} = \phi + \varphi \left(\ln P_{F,j,t} - \ln P_{N,j,t}\right) + \beta \left(\ln Y_{i,j,t} - \ln P_{j,t}\right) + X^{'}\theta + u_{i,j,t} \tag{A1}$$

where w is the share of the budget devoted to food, P is the true, unobserved price index for food (F), nonfood (N), and all goods, Y is household expenditure, X is a vector of household characteristics, and u is the error term. The true cost of living in year t, in place j, $P_{j,t}$, is a weighted average of the prices of food and non-food

$$\ln P_{j,t} = \alpha \ln P_{F,j,t} + (1-\alpha) \ln P_{N,j,t}$$

and those prices are measured with error (which is CPI bias) such that

$$\ln\left(P_{j,t}\right) = \ln\left(P_{j,0}\right) + \ln\left(1+\Pi_{j,t}\right) + \ln\left(1+E_{j,t}\right) \tag{A2}$$

where P_0 is the true price, Π is the CPI price, and E is the cumulative measurement error in the price index from year 0 to year t. Note that the measurement error would also apply to the prices of food and non-food in the same manner, and that aggregate error would also be a weighted function of the errors in food and non-food. Substituting (A2) into the Almost Ideal Demand System (A1) and rearranging terms yields

$$w_{i,j,t} = \phi + \varphi \left[\ln\left(1+\Pi_{F,j,t}\right) - \ln\left(1+\Pi_{N,j,t}\right)\right] + \beta \left[\ln Y_{i,j,t} - \ln\left(1+\Pi_{j,t}\right)\right] + X^{'}\theta + \varphi \left[\ln P_{F,j,0} - \ln P_{N,j,0}\right] - \beta \ln\left(P_{j,0}\right) + \varphi \left[\ln\left(1+E_{F,j,t}\right) - \ln\left(1+E_{N,j,t}\right)\right] - \beta \ln\left(1+E_{j,t}\right) + u_{i,j,t}$$

The functional form of estimating CPI bias (the Hamilton-Costa form)

$$w_{i,j,t} = \phi + \varphi \left[\ln\left(1+\Pi_{F,j,t}\right) - \ln\left(1+\Pi_{N,j,t}\right)\right] + \beta \left[\ln Y_{i,j,t} - \ln\left(1+\Pi_{j,t}\right)\right] + X^{'}\theta + \sum_{t=1}^{T} \delta_t D_t + \sum_{j=1} \delta_j D_j + u_{i,j,t}$$

follows directly and

$$\delta_t = \varphi \left[\ln\left(1+E_{F,j,t}\right) - \ln\left(1+E_{N,j,t}\right)\right] - \beta \ln\left(1+E_{j,t}\right)$$

is used to measure the extent of CPI bias. If we assume that the bias between food and non-food is constant and that both food and nonfood are equally biased it holds that

$$\ln\left(1+E_{j,t}\right) = \tfrac{-\delta}{\beta}$$

such that cumulative (percentage) CPI bias at t is

$$1 - \exp\left(\tfrac{-\delta}{\beta}\right)$$

For the 1888/1917 estimates of CPI bias, data on CPI measured relative price changes by region are unavailable. Therefore, estimates of relative price changes over time would collapse into the time dummy as there would be no regional variation. In this instance, the Engel cure used to estimate CPI bias is

$$w_{i,j,t} = \phi + \beta \left[\ln Y_{i,j,t} - \ln \left(1 + \Pi_{j,t} \right) \right] + X^{'}\theta + \sum_{t=1}^{T} \delta_t D_t + \sum_{j=1} \delta_j D_j + u_{i,j,t}$$

where now

$$\delta_t = \varphi \left[\ln \left(1 + \Pi_{F,j,t} \right) - \ln \left(1 + \Pi_{N,j,t} \right) \right] + \varphi \left[\ln \left(1 + E_{F,j,t} \right) - \ln \left(1 + E_{N,j,t} \right) \right] - \beta \ln \left(1 + E_{j,t} \right)$$

so that with the same assumptions as those above, and for a given value of φ and changes in relative prices, the cumulative CPI bias at t is

$$1 - \exp \left(\frac{\delta - \varphi [\ln(1+\Pi_{F,t}) - \ln(1+\Pi_{N,t})]}{-\beta} \right)$$

I use the estimate of φ from the 1917/1935 CPI bias regressions to estimate the 1888/1917 CPI bias, which is the same methodology adopted in Costa (2001). It is worth noting that this methodology does not require income to be deflated. If income were not deflated in the Engel curve then δ_t would be used to estimate the true cost of living rather than bias in that cost. The method does require estimates of the relative price changes over time, as these would have an effect on the demand for food generally.

3 Notes on the Theory of Household Economies of Scale

The traditional (Engel) estimates of economies of scale supposes that there is an economies of scale parameter θ such that n^{θ} would reflect the fact that for some goods the needs of the household do not increase exactly as much as the number of people.[40] For this reason, $0 \leq \theta \leq 1$. If $\theta = 1$ there would be no economies of scale, whereas if $\theta \in (0, 1)$ there would be economies of scale. One could therefore use $1 - \theta$ as a measure of household economies of scale. The lower the value of θ the greater the household scale economy for that particular good. We can think of integrating these economies of scale into the standard household cost function with

$$c(u, p, n) = n^{\theta} \alpha(p) u^{\beta(p)}$$

where the cost minimizing function takes family size as an argument. This gives us an indirect utility function

$$\ln u = \frac{\ln x - (\theta \ln n + \ln \alpha(p))}{\beta(p)}$$

If we use the standard specification that $\ln \alpha(p) = \sum \alpha_j \ln p_j$ and $\ln \beta(p) = \sum \beta_j \ln p_j$ the budget share equations (the Engel curve) becomes

$$w = \alpha + \beta \ln \left(\frac{x}{n} \right) + \beta (1 - \theta) \ln n$$

and we can compute the economy of scale parameter.

There are two problems with this specification. The fist problem is the identification problem noted by Pollack and Wales (1979). To see this, imaging modifying the household cost function to reflect the fact that additional people have larger or smaller effects that depend of the utility of the household (this is analogous to claiming that the economies of scale themselves are related to the welfare of the household, as the Barten model predicts). In this situation the cost function is now

$$c(u, p, n) = n^{\theta + \theta_1 \ln u} \alpha(p) u^{\beta(p)} = n^{\theta} \alpha(p) u^{\beta(p) + \theta_1 \ln n}$$

where the budget share equation is again

$$w = \alpha + \beta \ln \left(\frac{x}{n} \right) + \beta (1 - \theta) \ln n$$

but where now the indirect utility function is

[40] This discussion borrows heavily from Deaton (1997). The idea here is analogous to the standard $A + C^{\theta}$ formula used to derive adult equivalence, where A is the number of adults, C the number of children, and θ the adult-equivalent factor for children given their age and sex. Just as in the Engel case $0 \leq \theta \leq 1$.

$$\ln u = \frac{\ln x - (\theta \ln n + \ln \alpha(p))}{\beta(p) + \theta_1 \ln n}$$

Without some further appeals to substitutions between public and private goods, the Engel method give us no theoretical basis to call the estimate of θ a household scale economy. While the Barten model uses the notion of a scaling parameter, it is housed in a discussion of the income and substitution effects that changes in the prices of public and private goods would have on consumption. No such case is implied in the Engel methodology or procedure. This lack of theoretical depth has lead some to conclude that "the scaling [Engel] model, although identified on the data once specified, has the same empirical implications as other models in which the economies of scale are clearly different. As a result, whatever it is that the Engel method measures, there are no grounds for claiming that it is economies of scale" (Deaton 1997, p. 268).

The second problem is the restrictions that are placed on the demand function to yield estimates of economies of scale. From the demand equations given above we know that $\theta = 1 - \frac{\gamma}{\beta}$, where γ is the coefficient on the log of household size from the demand equation. For $0 \leq \theta \leq 1$ it must hold that $\frac{\gamma}{\beta} \in [0, 1]$, which implies that both (1) $|\gamma| \leq |\beta|$ and that (2) $sign(\gamma) = sign(\beta)$. If (1) does not hold the scale economy would be either be greater than 1 or less than 0 ($0 > \theta > 1$), and if (2) does not hold the scale economy would be greater than 1 ($\theta > 1$). While both of these conditions hold for food in almost all cases, it is not clear that they must hold for other goods. The Engel measure requires that the effect of household size on the budgetshare must be in the same direction as the effect of total expenditures, and the Barten model predicts the opposite for some private goods (such as food). Also, it is not clear why the effect on income must be larger than the effect of household size to compute Engel scale economy estimates. Comparing estimates of θ for various goods, then, is based upon a conjecture that does not give one firm grounds to make a comparison of economy of scale parameters when using the Engel procedure. Table A2 shows estimates of θ for the period 1888-1935 for each of the expenditure categories considered in the paper. The results imply that food away from home has more scale economies than food at home, certainly a questionable result. In addition to showing lack of a trend, the results for some expenditures, such a clothing, cannot be interpreted as household economies of scale because they fall outside of the bounds allowed for θ. These historical results confirm the dubious nature of Engel scale economy estimates.

On a related note, there is some theoretical justification for the linear forms for the Engel curves taken in this paper. Suppose that one wished to estimate a non-parametric Engel curve of the form

$$w = g(\ln x) + \varepsilon$$

where $g(*)$ is some non-parametric function of the log of expenditure, and where we would like to control for demographic effects z. One would like, assuming the demographic effects are linear, to estimate a partially linear model of the form

$$w = g(\ln x) + \varsigma z + \varepsilon$$

The problem is that the additive structure of the model above and the Slutsky symmetry conditions requires that $g(\ln x)$ be linear (see Blundell, Browning, and Crawford (2003) for a proof). Indeed, one method for having $g(*)$ not enter linearly would be to deflate $\ln x$ by a general household equivalence scale, which itself depends on estimates of household scale economies to the extent that additional persons in the household share public goods. This may be the reason why the first differencing results, where $g(\ln x)$ can take the value of any continuous function but where the household composition parameters enter linearly, mirror those of the linear specifications.

4 Data Appendix

4.1 The Consumer Expenditure Surveys

I used three consumer expenditure surveys in this paper, covering the years 1888-1890, 1917-1919, and 1935-1936. For the 1888-1890 survey, the sample was selected only from the following nine industries: pig iron, bar iron, steel, bituminous coal, coke, iron ore, cotton textile industry, wool textile industry and glass. Sample families, limiting to those representing more than two persons, were chosen from employer records. These households were then selected to provide detailed expenditure information to survey respondents, and in most instances expenditures were verified by local merchants. Twenty-four states were covered. In total, nearly 7,000 American families were surveyed. For more on the sampling see Logan (2006).

The 1917-1919 data were obtained from the surveys over 12,000 families of wage earners or "small salaried workers." As with the 1888-1890 survey, households were selected from employer records. Sample families were chosen such that there are only husband and wife families with at least one child; the salary earners had to earn than $2,000 per year; families had to reside in the community at least one year prior to the interview; families could not have more than three boarders; families could not be "slum" or "charity"; and non-English-speaking families had to reside in the United States for more than five years. All the selections are from ninety-nine cities throughout 42 states.

In the 1935-1936 survey, only native-born families living in the United States were selected. The sample covered 51 cities, 140 villages, and 66 farm counties throughout 30 states. Except for New York City, Columbus, OH, and the South, only white families were chosen. Families in large cities had to earn more than $500 a year and those in smaller localities had to earn more than $250 a year. There was no income limit on households in this survey, and self-employed households were included as well. The data used in this paper comes from a random sample of 6,000 families of the 61,000 who provided both income and expenditure information. Since the 1935-1936 survey was explicit in its desire to capture the expenditure of rural households, while the 1917-1919 and 1888-1890 selected almost exclusively on urban households, I used only the urban households from the 1935-1936 survey in the analysis, which is more than half of the 6,000 observations. For the estimates of CPI bias, I used the rural data as well, although due to missing values this added only 116 households from rural areas from the 1935-1936 survey. Table A3 shows the means of the expenditure categories and household size for each survey.

Construction of household expenditure for the three surveys was similar. While the construction of the clothing and food categories was straightforward, housing and entertainment vary somewhat by survey. Entertainment in 1888-1890 is comprised of expenditures on books, newspapers, vacations, and "other amusements." For 1917-1919 entertainment includes expenditures on movies, concerts, plays, lectures, dances, billiards, excursions, books, and "other amusements." For 1935-1936 entertainment includes movies, radios, sports clubs, social clubs, camping, fishing, hiking, sports (golf, baseball, horseback riding, tennis, etc.), bikes, skates, skis, billiards, boats, cameras, vacations, and "other recreational expenses." Housing in every survey year includes both rent and/or mortgage payments, lighting, and fuel expenditures. For 1917-1919 and 1935-1936, the list expands to include utilities such as electricity and sewage. I also constructed a housing expenditure variable that included expenditures on furniture and appliances, as these are likely goods to be used by and/or for multiple household members. The results in the text with regards to housing are robust to the inclusion of furniture and appliances, and other expenditures such as alcohol and tobacco prove to be consistent with the predictions for private goods in the Barten model (see Table A4).

4.2 Price Indices

Overall price change for 1888-1917 and overall and food price changes for 1917-1935 were calculated from the *Historical Statistics of the United States, Millennial Edition* (2006). For 1917-35, regional price indices were calculated using the *Handbook of Labor Statistics: 1950 Edition* (U.S. Bureau of the Census 1951) which contains price indices for 1917-1935 for food and all items for a sampling of cities in the United States. These were applied to the states from which the prices came, to construct a regional price index using the Census Bureau's regions (this gives four regions for the US). Assuming that the price index is a weighted sample of food and non-food, the price indices are used to create regional price indexes for food, non-food and all items for 1917-1935.

Table A1

The Effect of Household Size on Food and Clothing's Share of the Budget 1888-1935

	I	II	III
Method	1888	1917	1935
Linear (OLS)	0.045	-0.050	-0.016
	(10.25)	(15.11)	(1.25)
Fourier	0.047	-0.046	-0.024
	(10.75)	(13.65)	(1.89)
First Differencing	-0.026	0.020	0.003
	(6.85)	(7.17)	(0.23)
IV	0.057	-0.033	-0.003
	(12.56)	(9.53)	(0.24)
N	6809	12817	3534

The Effect of Household Size on Food in Food and Housing Expenditures 1888-1935

	I	II	III
Method	1888	1917	1935
Linear (OLS)	-0.017	0.048	-0.033
	(1.18)	(10.22)	(2.24)
Fourier	-0.017	0.047	-0.032
	(1.17)	(10.08)	(2.17)
First Differencing	0.001	-0.024	0.019
	(0.1)	(7.01)	(1.4)
IV	-0.092	0.053	-0.048
	(6.09)	(10.93)	(10.59)
N	6809	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household. The dependent variable in each regression is the share of household expenditure devoted to food and clothing (top panel) and the share of expenditure devoted to food from food and housing expenditure (bottom panel). Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.

Table A2

Engel Estinmates of Household Economies of Scale, 1888-1935

	I	II	III
	1888	1917	1935
Food	0.809	0.426	0.709
Food at Home		0.448	0.520
Food Away from Home		0.313	0.378
Clothing	-2.851	-0.0005	-1.956
Entertainment	0.714	0.398	0.597
Housing	0.708	0.002	19.038
N	6809	12817	3534

Note: Estimates are from OLS regressions reported in Tables 1-5, where the budgetshare is regressed on per capita expenditure, household characteristics, geographic location, and industry. The Engel measure of household scale economies is calculated as the coefficient on househld size divided by the coefficient estimate on per capita expenditure subtracted from unity.
Low values imply large household economies of scale, since the value here is the power by which the effective household size increases for that consumption good with an additional member of the household.

Table A3

Summary Statistics from Historical Household Surveys, 1888-1935

	I	II	III
Variable	1888	1917	1935
Household Size	4.7	4.9	3.7
	(2.11)	(1.64)	(1.45)
Food Share	44.5%	39.2%	38.9%
	(.089)	(.079)	(.093)
Clothing Share	16.7%	16.2%	10.9%
	(.065)	(.050)	(.056)
Entertainment Share	1.9%	3.2%	3.5%
	(.024)	(.009)	(.027)
Housing Share	13.7%	13.6%	14.3%
	(.081)	(.069)	(.088)
N	6809	12817	3534

Note: Estimates are the mean values, based on Author's calculation.
Standard Deviations are listed in parentheses.

Table A4

The Effect of Household Size on Housing and Furniture's Share of the Budget 1888-1935

Method	I 1888	II 1917	III 1935
Linear (OLS)	-1.476	-0.053	-0.026
	(3.41)	(15.01)	(3.04)
Fourier	-1.502	-0.052	-0.022
	(3.47)	(14.47)	(2.56)
First Differencing	-1.673	-0.052	-0.019
	(3.88)	(14.66)	(2.14)
IV	0.072	-0.056	-0.033
	(1.6)	(15.19)	(3.67)
N	6809	12817	3534

The Effect of Household Size on Alcohol and Tobacco's Share of the Budget 1888-1935

Method	I 1888	II 1917	III 1935
Linear (OLS)	0.005	0.002	0.008
	(1.19)	(1.1)	(1.14)
Fourier	0.006	0.002	0.008
	(1.47)	(1.15)	(1.18)
First Differencing	-0.004	0.000	-0.001
	(1.37)	(0.33)	(0.11)
IV	0.008	0.002	0.007
	(2.08)	(1.11)	(0.88)
N	6809	12817	3534

Note:
Each entry is the coefficient estimate for the log of household size for that method, year pairing.
Each entry comes from a separate regression that includes the fraction of the household employed, the state of residence, the industry of the household head, and demographic shares of the household. The dependent variable in each regression is the share of household expenditure devoted to housing, furniture, and appliances (top panel) and share of expenditure devoted to alchohol and tobacco (bottom panel). Robust t-statistics are listed under coefficient estimates in parenthses.

See section 3 of the text for details on the estimation procedure.
See the data appendix for the expenditure categories that comprise housing expenditures.