ORIGINAL PAPER

Family Size and Food Consumption: Exploring Economies of Scale with Panel Data

José M. Casado[1] · José Ignacio Giménez-Nadal[2] · José M. Labeaga[3] · José Alberto Molina[2]

Abstract

We analyze the effects of family size on food consumption. Specifically, we explore the economies of scale that emerge in families, taking advantage of a very rich panel data for Spanish consumers. We hypothesize that family public goods induce economies of scale in consumption and, consequently, larger families should have higher per capita consumption of private goods, such as food. To that end, we estimate the Quadratic Almost Ideal Demand System to derive own-price and income elasticities, which allows us to test the economies of scale conditions on food consumption. Our results show that the positive relationship between per capita food expenditure and family size holds in only 585 observations, with these families being characterized as having per capita income in the two highest centiles of the distribution. Our results should be interpreted in the context of the particular economic crises of the 1970s and 1980s, as a consequence of the economic shocks due to very high oil prices that affected Spain.

Keywords Family size · Food consumption · Economies of scale · Panel data · Unobserved heterogeneity · Demand systems

JEL Classification C12 · C23

Introduction

The relationship between family size and food consumption has been studied in the literature, using cross-sectional data. Specifically, consumption of food is a very relevant analysis, given that it is usually employed to measure living standards of families (poverty and inequality comparisons) and, consequently, the relationship is very important from a policy perspective.

The theoretical foundations of the relationship between family size and food consumption were initially provided by Barten (1964), whose model was subsequently employed by Deaton and Paxson (1998). The Barten (1964) model assumes that families consume both public and private goods and that the most plausible source of economies of scale is the presence of public goods that can be shared within the family (e.g., housing or fuel). In the presence of public goods, a couple family is better-off than a single family with the same per capita income or expenditure, because the resources made available by sharing family public goods can be used to acquire a greater quantity of both private and public goods. When there are family scale economies associated with public goods, larger families should have comparatively higher per capita private consumption, such as food, in comparison to smaller families. However, this theoretical prediction has not been confirmed by the empirical literature. In fact, much of the evidence is persistent in confirming the opposite; i.e., food expenditure per capita falls as the number of family members rises (the Deaton and Paxson puzzle).

Several authors have explored the relationship between family size and per capita food consumption using cross-section data. Gan and Vernon (2003) postulated the possibility that food is more public than other private goods (e.g., returns to scale in food consumption, particularly in food preparation). In this sense, the food share increases with family size, relative to more public goods and decreases

✉ José Alberto Molina
jamolina@unizar.es

1 Bank of Spain, Madrid, Spain

2 University of Zaragoza and IEDIS, Saragossa, Spain

3 UNED, Madrid, Spain

relative to more private goods. However, these authors provide no evidence of greater economies of scale in food than in clothing or transportation. Crossley and Lu (2018) develop a slight modification of Barten's model, adding home production of food and two types of food that differ in their preparation time. The authors illustrate the case in which economies of scale in food preparation is plausible, but the effect is difficult to identify due to data limitations. Other unsuccessful explanations of the paradox between the theory and the evidence are bulk discounts (Abdulai, 2003; Gibson & Kim, 2018; Griffith et al., 2009), errors in food expenditure data correlated with family size (Brzozowski et al., 2016; Gibson, 2002; Gibson & Kim, 2007), the possibility of different wastage schemes depending on size for different families, the effects of the price elasticity of food, and the effects of the quality of food (Deaton & Paxson, 1998).

Given that prior evidence using cross-sectional data is far from reconciled with the theory through applied analysis, we explore the attractiveness of a rich panel data on consumption (a family can be observed up to a maximum of 27 consecutive quarters) to analyze the relationship between family size and food consumption. We base our work on the Barten Model and estimate the Quadratic Almost Ideal Demand System (QUAIDS) to derive own-price and income elasticities. The attractiveness of our Spanish panel data on consumption comes from the availability of quarterly consumption data, which allows us to explore two additional methodological channels to reconcile theory and evidence that have not been previously studied in depth.

First, we observe that the control of family unobserved heterogeneity (time differences in tastes for food across families) does not provide evidence according to the theory. Second, we employ our long-time series of family data to estimate a Quadratic Almost Ideal Demand System model (QUAIDS, Banks et al., 1997). The income and price elasticities allow us to test the underlying assumption of the economies of scale conditions derived by Deaton and Paxson (1998, 2003); that is to say, the extent to which the own price food elasticity is less than the income elasticity. We test this by assuming the existence of two goods, i.e., food and housing, but in the context of several commodities. The estimation of the QUAIDS allows us to find estimated elasticities agreeing with the theory for 585 observations, with these being characterized as having per capita income in the two highest centiles of the distribution.

The remainder of the paper is organized as follows. Sect. "Theoretical Model of Family Size and Consumption" describes Barten's model. Sect. "Data and Baseline Estimations" describes the panel data and shows baseline estimations. Sect. "Explorations of the Relationship Between Family Size and Food Consumption Using Panel Data" presents possible explanations of the paradox by using our panel data, i.e., applying single-equation fixed effects

estimators to the Engel curves and estimating complete demand systems. Finally, Sect. "Conclusions" presents our main conclusions.

Theoretical Model of Family Size and Consumption

Barten's model (Barten, 1964), used by Deaton and Paxson (1998), assumes that families consume both public and private goods, and that economies of scale are generated by the presence of public goods that can be shared within the family. According to this model, a family with n adult members allocates its total expenditure x across two goods, a private (e.g., food f) and a public good (e.g., housing h), optimizing the following problem:

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

$$s.t. p_f\left(\frac{q_f}{n}\right) + p_h\left(\frac{q_h}{n}\right) = \frac{x}{n} \tag{1}$$

where $\phi_f(n)$ and $\phi_h(n)$ are the scaling functions showing some commodity-specific scale-economies so that effective family size for the consumption of each good is not n but rather $\phi_i(n)$, $i = f, h$. For pure private (f) and pure public (h) goods, the utility function is $nv\left(\frac{q_f}{n}, q_h\right)$. Solving (1), the per capita food demand function is

$$\frac{q_f}{n} = \frac{\phi_f(n)}{n} g_f\left(\frac{x}{n}, \frac{p_f \phi_f(n)}{n}, \frac{p_h \phi_h(n)}{n}\right) \tag{2}$$

Taking logs and differentiating with respect to $\ln n$ we get the condition for per capita food consumption to increase with family size, holding constant the per capita expenditure:

$$\Upsilon^* = \frac{\partial \ln(p_f q_f/n)}{\partial n} = \sigma_h(e_{fx} + e_{ff}) - \sigma_f(1 + e_{ff}) \tag{3}$$

In Eq. (3), e_{ff} and e_{fx} are the own-price and income elasticities of food; $\sigma_i = 1 - \frac{\partial \ln \phi_i(n)}{\partial \ln n}$ for $i = f$, and h is the commodity-specific (technological) economy of scale measure. A pure private good has $\sigma_i = 0$ and a pure public good has $\sigma_i = 1$. Υ^* is the elasticity of per capita food consumption with respect to family size, and there are economies of scale when the condition given in (3) is greater than zero:

$$\Upsilon^* = \sigma_h(e_{fx} + e_{ff}) - \sigma_f(1 + e_{ff}) > 0 \tag{4}$$

At constant per capita expenditure, per capita food consumption increases with family size when (1) food has limited substitutes, i.e., e_{ff} is small in absolute value and lower

than e_{fx}; (2) food has significantly less economies of scale than housing, i.e., σ_f/σ_h is small. In other words, in the basket of goods of a family, there are private commodities like food with low own- and cross-price elasticities where the income effect dominates, and thus the per capita consumption of this good should increase with family size.

Even under the assumption that since there are few substitutes for food, cross-goods interrelationships could hide some effect that would solve the puzzle between the theory and the evidence and, consequently, Horowitz (2002) and Deaton and Paxson (2003) reformulate the previous expression to take into account the situation when the number of commodities is more than two and there are substitution and complementary effects among them, and between food and the other goods in the context of demand systems:

$$\bar{\sigma}\left(e_{fx} + e_{ff}\right) - \sigma_f\left(1 + e_{ff}\right) < \sum_{k \neq f} \bar{e}_{fk}\left(\sigma_k - \bar{\sigma}\right) \tag{5}$$

where $\bar{e}_{fk}$ is the compensated elasticity of the demand for food with respect to the price of good k, and $\bar{\sigma}$ is the budget share weighted average of the economies of scale parameters for goods in the system, except food.

Data and Baseline Estimations

Data

We use the Spanish Permanent Consumption Survey (EPC), from the second quarter of 1977 to the fourth quarter of 1983, carried out by the Spanish National Statistics Institute (INE). Despite that it is an old dataset, it is very useful in the current context as it covers a high time dimension, as a family can be observed up to a maximum of 27 consecutive quarters. The survey contains expenditure information for 130 items and an important range of socio-demographic variables. To the best of our knowledge, this is the longest panel on family consumption expenditures, covering a wide range of commodity groups. Almost two-thirds of the families remain in the survey for at least 3 years (12 quarters), and this allows us not only the estimation of models with the possibility to control for unobserved heterogeneity, but also to follow families experiencing changes in their structure (e.g., family size).

We select families with fewer than 7 missing quarters on food expenditure, using the largest sample period for each family. We omit families with no food expenditure information. We restrict the sample to families observed for at least 12 quarters, without missing values for the relevant variables, so that we have changes in family size for some families. This sample selection leaves us with an unbalanced panel (Tables 1 and 2) of 1,452 families and a total of 28,926 observations.

Baseline Estimations

We estimate Engel curves using both a non-parametric and parametric approaches. For the non-parametric approach, we closely follow the empirical strategy of Deaton and Paxson (1998) and we fit Engel curves for families of different sizes and composition, estimating the following equation:

$$\int E\left(w_f | i, z\right) g(z) dz \tag{6}$$

where w_f is the share of food consumption, i is an index describing the composition of the family, z is the log of per capita expenditure, and $g(z)$ is a nonparametric kernel estimate of density, so (6) adjusts food share averages for each type of family, weighted by density of per capita expenditure.

Figure 1 shows kernel estimates of food shares for 1-adult and 2-adult families with no children. Observations are weighted by density of the per capita expenditure. We include confidence intervals (defined at the 95% level) to see if the differences between the two kernel distributions are statistically significant. We observe that in most values of the total per capita expenditure, the average food share is larger in smaller families. Thus, when we analyze the relationship between family size and food shares by comparing 1-adult and 2-adult families at constant per capita expenditure, we see that the food share (and therefore the per capita expenditure on food) declines with family size, which confirms the Deaton-Paxson puzzle. Figure 2 show similar patterns, although we now consider more types of families. When we consider childless families from 1 to 4 adults, we observe that as the number of adults increases, at constant per capita expenditure, the average food share decreases, especially between 1-adult, 2-adult, and 3-adult families. This also confirms the Deaton-Paxson puzzle.

Table 3 provides average food shares (and standard errors) for different types of families. In the first set of families, we compare families with no children, and as the number of adults increases, the food share decreases. Similar patterns emerge when we compare one adult-one child families versus two adults-two child families, as the average of food share decreases from 0.50 to 0.47. When we compare one adult-two child families versus two adults-four child families, the average food share decreases from 0.59 to 0.51, respectively. This points to larger families having lower per capita food expenditure, indicating that, at constant per capita expenditure, the food share, and, therefore, the per

Table 1 Number of families per period remaining in the sample

Number of quarters	Full data set			Final sample		
	Observations	Number of families	%Families	Observations	Number of families	%Families
1	169	169	5.47	0	0	0.00
2	194	97	3.14	0	0	0.00
3	468	156	5.05	0	0	0.00
4	612	153	4.95	0	0	0.00
5	345	69	2.23	0	0	0.00
6	378	63	2.04	0	0	0.00
7	686	98	3.17	0	0	0.00
8	720	90	2.91	0	0	0.00
9	423	47	1.52	0	0	0.00
10	340	34	1.10	0	0	0.00
11	682	62	2.01	0	0	0.00
12	696	58	1.88	619	52	3.58
13	533	41	1.33	2786	211	14.53
14	532	38	1.23	1245	89	6.13
15	915	61	1.97	869	59	4.06
16	1152	72	2.33	831	52	3.58
17	612	36	1.17	3375	199	13.71
18	684	38	1.23	627	34	2.34
19	1330	70	2.27	759	40	2.75
20	1420	71	2.30	939	47	3.24
21	966	46	1.49	2161	103	7.09
22	792	36	1.17	546	25	1.72
23	2116	92	2.98	1304	57	3.93
24	2184	91	2.95	671	28	1.93
25	450	18	0.58	896	36	2.48
26	546	21	0.68	650	25	1.72
27	34,074	1262	40.85	10,648	395	27.20
Total	54,019	3,089	100	28,926	1452	100

Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983, and we select families reporting full information for at least 12 consecutive quarters (3 years)

capita expenditure on food, declines with family size, thus confirming the Deaton-Paxson puzzle.

We also apply a parametric approach for the estimation of Engel curves. To that end, we follow the specification proposed by Deaton and Paxson (1998), as follows:

$$w_{hf} = \beta \ln \frac{x_i}{n_i} + \gamma \ln n_i + \sum_{k=1}^{K-1} \eta_{ik} \frac{n_{ik}}{n_i} + \zeta V_i + \pi_t + \varepsilon_i \qquad (7)$$

where w_{if} is the food share of total consumption by family i, $\frac{x_i}{n_i}$ is per capita expenditure with n_i being the size of family "i", and $\frac{n_{ik}}{n_i}$ are the ratios to family size of the number of males and females of different ages (ratio of male children aged 0–5, 6–11 and 12–17, ratio of female children aged 0–5, 6–11 and 12–17, the ratio of males aged 18–64 and the ratio of females aged 18–64). V is a vector of socio-demographic controls, including dummy variables to control for whether the family head (FH) works in an agricultural sector, is a blue collar worker, has primary education, FH has tertiary education, owns the home, and lives in a rural area (see Table 4 for summary statistics of the socio-demographic characteristics of the households). We also include aggregate effects common to all families through annual and quarterly dummies.

We estimate Eq. (7) using Ordinary Least Squares (OLS) and Instrumental Variables (IV), since consumption (on food and total consumption) could be measured with errors and thus correlated with unobservable characteristics in (6). We follow Gibson (2002) and use proxies for income as instruments for total expenditure. We use the average number of school years of all adults in the family, the age of the head of family, a dummy variable that indicates whether the family

Table 2 Number of families per quarter

Number of quarters	Full data set Number of families	Final sample Number of families	%Families covered
77/II	1993	818	41.04
77/III	2001	831	41.53
77/IV	2001	866	43.28
78/I	2001	899	44.93
78/II	2001	955	47.73
78/III	2001	968	48.38
78/IV	2001	1099	54.92
79/I	2001	1130	56.47
79/II	2001	1158	57.87
79/III	2001	1169	58.42
79/IV	2001	1270	63.47
80/I	2001	1299	64.92
80/II	2001	1299	64.92
80/III	2001	1175	58.72
80/IV	2001	1310	65.47
81/I	2001	1318	65.87
81/II	2001	1298	64.87
81/III	2001	1139	56.92
81/IV	2001	1150	57.47
82/I	2001	1121	56.02
82/II	2001	1098	54.87
82/III	2001	990	49.48
82/IV	2001	990	49.48
83/I	2001	938	46.88
83/II	2001	926	46.28
83/III	2001	856	42.78
83/IV	2001	856	42.78
Total	54,019	28,926	

Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983, and we select families reporting full information for at least 12 consecutive quarters (3 years)

Fig. 1 Non-parametric Engel curves, 1-adult and 2-adult families with no children. Data comes from the Spanish Permanent Survey of Consumption EPC), years 1977 to 1983. Food shares are defined as the expenditure on food out of total expenditure. Observations are weighted by density of the per capita expenditure

Fig. 2 Non-parametric Engel curves by number of adults, families with no children. Data comes from the Spanish Permanent Survey of Consumption EPC), years 1977 to 1983. Food shares are defined as the expenditure on food out of total expenditure. Observations are weighted by density of the per capita expenditure

has a second home, and the per capita expenditure of the previous year.

Columns 1 and 2 of Table 5 show the results of estimating Eq. (7) with both OLS and IV for the whole sample of families. Family size appears to exert a negative and statistically significant effect on the food budget share, holding per capita expenditure constant. An increase of 1 per cent in the logarithm of fam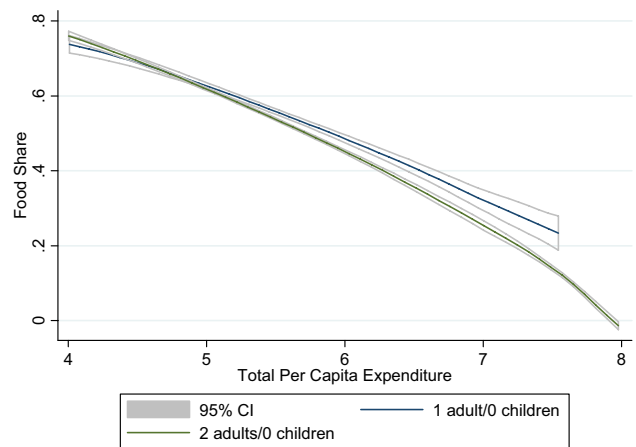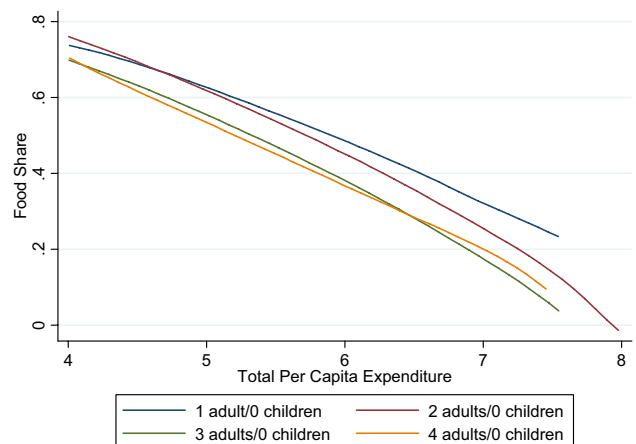ily size decreases the food budget share by 6 per cent. IV estimates show the same pattern as OLS estimates. Over- and under-identification tests indicate the adequacy of the instruments. Thus, these results point to the Deaton-Paxson puzzle. We estimate Eq. (6) for a sample of families with more than one adult (Columns 3 and 4 of Table 5) and with a sample of families where the head is employed (trying to control for separability between consumption of food and labor supply), and we observe very similar results, since the puzzle is reflected in the results.

Table 3 Average food share for each type of family

Adults-children	Food in total expenditures
1–0	0.587
	(0.006)
2–0	0.574
	(0.003)
3–0	0.514
	(0.004)
4–0	0.504
	(0.005)
5–0	0.481
	(0.009)
1–1	0.500
	(0.016)
2–2	0.476
	(0.003)
X	
1–2	0.591
	(0.022)
2–4	0.511
	(0.007)
2–1	0.521
	(0.005)
4–2	0.517
	(0.007)

Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983. Food shares are defined as the expenditure on food as a proportion of total expenditure. Shares are weighted averages using weights obtained non-parametrically using kernel estimation of per capita total family expenditure

Table 4 Summary statistics of explanatory variables

	Mean	SD
ln (family size)	1.180658	(0.499)
ln (per capita expenditure)	5.273935	(0.577)
Ratio male children 0–5	0.0207035	(0.072)
Ratio female children 0–5	0.0200773	(0.071)
Ratio male children 6–11	0.0410748	(0.097)
Ratio female children 6–11	0.0402878	(0.097)
Ratio male children 12–17	0.0548593	(0.110)
Ratio female children 12–17	0.044838	(0.101)
Ratio males 18–64	0.2784027	(0.203)
Ratio females 18–64	0.3047939	(0.213)
Head works in agricultural sector	0.1382274	(0.345)
Head is blue collar worker	0.4592172	(0.498)
Head with primary education	0.234044	(0.423)
Head with tertiary education	0.3448783	(0.475)
Home ownership	0.2318868	(0.422)
HH in rural area	0.581134	(0.493)
Observations	28,926	

Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983, and we select families reporting full information for at least 12 consecutive quarters (3 years)

Explorations of the Relationship Between Family Size and Food Consumption Using Panel Data

In this section, we explore the panel structure of our data to analyze the relationship between family size and food consumption. Panel data allows us to consider the issue of the unobserved heterogeneity of families and, additionally, to estimate demand systems that allow us to derive income and own-price elasticities for testing Eq. (4).

The Role of Unobserved Heterogeneity

Deaton and Paxson (1998) acknowledge that larger families may have different tastes for food. If tastes for food differ in larger families, this could explain the puzzle. However, none of the prior studies have been able to consider differences in tastes for food across families, given that those studies use a cross-section of families and thus cannot control for the unobserved heterogeneity of families. Given that we have a panel of families, we can apply panel data models to parametrically estimate the Engel curve. To that end, the previous parametric specification of the Engel equation (Eq. 6) is improved to allow for time-invariant unobserved family heterogeneity.

We determine this over time, considering the unobserved family heterogeneity if food expenditure per capita falls as the family size rises, when per capita expenditure is held constant. To that end, we use the long-time panel of families (T = 27), and estimate the following augmented Engel curve:

$$w_{iht} = \alpha_i + \beta_i \ln \frac{x_{ht}}{n_{ht}} + \gamma_i \ln n_{ht} + \sum_{k=1}^{K-1} \eta_{htk} \frac{n_{htk}}{n_i} + \zeta V_{ht} + \pi_t + \vartheta_h + \varepsilon_{iht} \tag{8}$$

where i is the commodity, h is the family and t the period. ϑ_h is the individual family unobserved component. In the case of food shares, we estimate (8) for total food, and for food at home (ingredients) and food away from home.

Table 6 shows the results of estimating (8) for the full sample of families. Regarding the results for all food, we observe that the coefficient for family size becomes negative—but non-statistically significant—which is also applicable to the group of food at home. Regarding food away from home, the coefficient is positive and statistically

Table 5 Estimates of the food Engel curve

Food share	All families		Families with more than one adult		Employed head of families	
	OLS	IV	OLS	IV	OLS	IV
ln (families size)	− 0.062***	− 0.052***	− 0.081***	− 0.072***	− 0.075***	− 0.062***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)
ln (per capita expenditure)	− 0.144***	− 0.116***	− 0.148***	− 0.120***	− 0.148***	− 0.112***
	(0.002)	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)
Ratio male children 0–5	− 0.099***	− 0.106***	− 0.080***	− 0.085***	− 0.053***	− 0.048**
	(0.015)	(0.018)	(0.015)	(0.018)	(0.016)	(0.020)
Ratio female children 0–5	− 0.102***	− 0.104***	− 0.094***	− 0.097***	− 0.080***	− 0.077***
	(0.015)	(0.018)	(0.015)	(0.018)	(0.016)	(0.020)
Ratio male children 6–11	− 0.042***	− 0.062***	− 0.024**	− 0.043***	− 0.005	− 0.012
	(0.010)	(0.011)	(0.010)	(0.012)	(0.012)	(0.014)
Ratio female children 6–11	− 0.072***	− 0.070***	− 0.064***	− 0.065***	− 0.037***	− 0.023
	(0.010)	(0.012)	(0.010)	(0.012)	(0.012)	(0.014)
Ratio male children 12–17	− 0.031***	− 0.033***	− 0.004	− 0.006	0.035***	0.040***
	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)	(0.013)
Ratio female children 12–17	− 0.090***	− 0.100***	− 0.070***	− 0.081***	− 0.052***	− 0.057***
	(0.009)	(0.010)	(0.009)	(0.011)	(0.011)	(0.013)
Ratio males 18–64	− 0.061***	− 0.078***	− 0.062***	− 0.080***	− 0.034***	− 0.037***
	(0.005)	(0.006)	(0.005)	(0.006)	(0.008)	(0.010)
Ratio females 18–64	− 0.047***	− 0.054***	− 0.049***	− 0.060***	− 0.014*	− 0.022**
	(0.005)	(0.005)	(0.005)	(0.006)	(0.009)	(0.010)
Head works in agricultural sector	0.032***	0.035***	0.033***	0.037***	0.045***	0.054***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Head is blue collar worker	0.007***	0.008***	0.007***	0.007***	0.013***	0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Head with primary education	0.034***	0.052***	0.037***	0.055***	0.033***	0.052***
	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)
Head with tertiary education	− 0.026***	− 0.036***	− 0.030***	− 0.040***	− 0.034***	− 0.046***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Home ownership	− 0.005**	− 0.006**	− 0.003	− 0.004	0.001	0.002
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
HH in rural area	0.013***	0.014***	0.009***	0.009***	(0.002)	(0.005)
	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)
Observations	28,926	22,624	26,773	20,815	19,629	14,519
R-squared	0.347	0.338	0.350	0.340	0.331	0.315
F-test	305.3	172.4	290.0	159.1	198.5	100.6

Robust standard errors in parenthesis. Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983, and we select families reporting full information for at least 12 consecutive quarters (3 years). Food share is defined as the expenditure on food as a proportion of total expenditure. The instrumental variables are the average number of school years of each adult in the family, the age of the head of family, a dummy variable that indicates whether the family has a second home, and the per capita expenditure of the previous year. The last two columns show the estimation results with a sample consisting of the families with more than one adult where the head of the family is full-time employed as wage earner with a permanent job, implicitly assuming separability between the consumption of goods and labor supply). Dummy variables for the 17 regions of residence, and year and quarter dummies are also included. *Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level

significant. Thus, despite that the results change for the groups of food, and food at home, these results do not help to explain the puzzle, as the theoretical prediction implies that the coefficients should be positive and statistically significant. This conclusion is clearer when we consider the

sample of families with more than one adult. The GMM panel data estimates show that the coefficient for family size becomes negative and statistically significant, which is also applicable to the group of food at home. For the group of food away from home, the coefficient is positive

Table 6 Coefficients of the log(family size) on budget shares, panel data GMM estimation

	All families			Families with more than one adult		
	Pool		Panel data	Pool		Panel data
	OLS	IV	GMM	OLS	IV	GMM
Food	− 0.062***	− 0.052***	− 0.023	− 0.081***	− 0.072***	− 0.052*
	(0.003)	(0.003)	(0.024)	(0.003)	(0.004)	(0.027)
Food at home	− 0.066***	− 0.056***	0.035	− 0.085***	− 0.076***	− 0.066**
	(0.003)	(0.003)	(0.023)	(0.004)	(0.004)	(0.026)
Food away from home	0.004***	0.004***	0.013***	0.004***	0.004***	0.013***
	(0.001)	(0.000)	(0.004)	(0.001)	(0.001)	(0.005)
Alcohol & Tobacco	0.005***	0.005***	0.012**	− 0.003***	− 0.003**	0.009
	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.006)
Clothing	0.007***	0.010***	0.008	0.012***	0.016***	0.003
	(0.001)	(0.002)	(0.011)	(0.002)	(0.002)	(0.012)
Housing	− 0.036***	− 0.043***	− 0.083***	− 0.029***	− 0.037***	− 0.066***
	(0.002)	(0.002)	(0.015)	(0.002)	(0.002)	(0.017)
Services	0.033***	0.041***	0.041***	0.031***	0.042***	0.028*
	(0.002)	(0.002)	(0.014)	(0.002)	(0.003)	(0.016)
Leisure	0.021***	0.023***	0.033***	0.026***	0.029***	0.039***
	(0.001)	(0.001)	(0.005)	(0.001)	(0.001)	(0.006)
Observations	28,926	22,624	22,624	26,773	20,927	20,925
Number of families			1457			1365

Robust standard errors in parenthesis. Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983, and we select families reporting full information for at least 12 consecutive quarters (3 years). Budget shares are defined as the expenditure on the reference group as a proportion of total expenditure. Regressions also include the logarithm of per capita expenditure, the ratio to family size of the number of male and females aged 0–5, 6–11, 12–17 and 18–64, a dummy variable that indicates whether the head of family works in an agricultural sector, family is blue collar, primary education or tertiary education ref.: secondary education) of family head, home ownership, family lives in a rural area, dummies for the 17 regions of residence, and year and quarter dummies. The instrumental variables used in both the IV and GMM panel data estimation are the average number of school years of each adult in the family, the age of the head of family, a dummy variable that indicates whether the family has a second home, and the per capita expenditure of the previous year. *Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level

and statistically significant. The latter results are in line with the results shown in Table 5, indicating that consideration of the family unobserved heterogeneity (non-time-varying differences in tastes for food) in the estimation of the Engel equation does not help to resolve the paradox between theory and evidence.

The Estimation of Own-Price and Income Elasticities with a Demand System

According to Barten's model, the condition for per capita food consumption to increase with family size, holding constant the per capita expenditure, is captured by: $\sigma_h(e_{fx} + e_{ff}) - \sigma_f(1 + e_{ff}) > 0$, where e_{ff} and e_{fx} are the own-price and income elasticities of food, and $\sigma_i = 1 - \frac{\partial \ln \phi_f(n)}{\partial \ln n}$ for $i = f$, and h is the commodity-specific (technological) economy of scale measure. Given that a pure private good has $\sigma_i = 0$ and a pure public good has $\sigma_i = 1$, and if food (or

ingredients) has low own- and cross-price elasticities, we should expect that the income effect dominates, and thus the per capita consumption of this good should increase with family size.

The fact that we count with a panel of families observed during a long period allows us to estimate demand systems, since prices vary so as to allow identification of their effects. We propose to estimate a flexible system in terms of income and price responses: the QUAIDS (Quadratic Almost Ideal System, Banks, see Banks et al., 1997) from which we derive price and income elasticities.

The QUAIDS system is often used in the literature to model consumer demand with family data and is based on price-independent generalized logarithmic (PIGLOG) preferences, with Engel curves that are modelled as budget shares being a quadratic function of the log–budget. It has the advantage of a flexible underlying utility function and allows imposing the restrictions of a consistent demand system, like homogeneity and symmetry. For each i = 1,2, …, N goods

and the corresponding budget shares w_i, QUAIDS forms the following non-linear system of equations:

$$w_{iht} = \alpha_{ih} + \sum_j \gamma_{ij} lnp_{jt} + \beta_i ln\left[\frac{x_{ht}}{a(p_t)}\right] + \frac{\lambda_i}{b(p_t)}\left\{ln\left[\frac{x_{ht}}{a(p_t)}\right]\right\}^2 + v_{it} \tag{9}$$

for i = 1, 2, …, N goods and j = 1, …, N with consumption budget x_{ht} for family h and period t and prices p_{it}. Price indices are:

$$a(p_t) = \alpha_0 + \sum \alpha_{ih} lnp_{it} + \frac{1}{2}\sum\sum\gamma_{ij} lnp_{it} lnp_{jt} \tag{10}$$

$$b(p_t) = \prod p_{it}^{\beta_i} \tag{11}$$

Further explanatory variables that account for taste shifts in family consumption, such as demographic characteristics, are added to α_{ih}, so we include a linear specification $\alpha_{ih} = \alpha_{ih}(V_{ht})$, where V_{ht} contains the same variables included in (7). We decompose v_{it} into a fixed effect and a mixed error.

To calculate income and price elasticities, we derive (8) with respect to lnx_h and lnp_j (we omit in the expression of elasticities subindex t) to obtain:

$$\mu_{ih} = \frac{\partial w_{ih}}{\partial lnx_h} = \beta_i + \frac{2\lambda_i}{b(p)}\left\{ln\left[\frac{x_{ht}}{a(p)}\right]\right\}$$

$$\mu_{ijh} = \frac{\partial w_{ih}}{\partial lnp_j} = \gamma_{ij} - \mu_{ih}\left(\alpha_{ih} + \sum_k \gamma_{jk} lnp_k\right) - \frac{\lambda_i\beta_j}{b(p)}\left\{ln\left[\frac{x_{ht}}{a(p)}\right]\right\}^2$$

Then, income elasticities are given by $e_{iih} = \frac{\mu_{ih}}{w_{ih}} + 1$ and uncompensated price elasticities by $e_{ijh}^U = \frac{\mu_{ijh}}{w_{ih}} - \delta_{ij}$, with $\delta_{ij} = 1$ if i = j and $\delta_{ij} = 0$ if i ≠ j. Using the Slutsky conditions, we express the compensated price elasticities as $e_{ijh}^C = e_{ijh}^U + e_{iih}w_{jh}$. With figures of elasticities at hand, we can test whether (4) holds, assuming hypothetical values for the commodity-specific (technological) economy of scale measure for the six commodities of our demand system (e.g., food, alcohol and tobacco, clothing, housing, services, and other goods). Since both price and income elasticities are family-specific, in the case of the condition being satisfied, we can identify the sample of families where it holds.

First, we use average values of the income and own-price elasticities of food (Table 7) to establish whether the condition holds. When we only consider the income and price elasticities of food and we assume $\sigma_f = 0$ and $\sigma_h = 1$, the difference $\sigma_h(e_{fx} + e_{ff}) - \sigma_f(1 + e_{ff})$ is significantly equal to

Table 7 Estimated demand elasticities, QUAIDS model

	Food	Alcohol & Tobacco	Clothing	Housing	Services	Others
Compensated						
Food	− 0.634	0.077	0.112	0.103	0.183	0.160
Alcohol & Tobacco	1.158	− 1.094	− 0.270	− 0.505	− 0.194	0.905
Clothing	0.264	− 0.185	− 0.310	0.275	0.494	− 0.537
Housing	0.531	− 0.297	0.280	− 0.444	0.455	− 0.525
Services	0.448	− 0.101	0.356	0.354	− 1.701	0.645
Others	0.801	0.373	− 0.416	− 0.318	0.128	− 0.568
Uncompensated						
Food	− 0.957	0.046	0.050	0.052	0.112	0.044
Alcohol & Tobacco	0.856	− 1.124	− 0.328	− 0.553	− 0.261	0.796
Clothing	− 0.167	− 0.227	− 0.393	0.206	0.398	− 0.692
Housing	− 0.079	− 0.356	0.164	− 0.541	0.320	− 0.744
Services	− 0.167	− 0.161	0.238	0.257	− 1.837	0.424
Others	− 0.125	0.283	− 0.592	− 0.465	− 0.077	− 0.901
Budget						
Food	0.654					
Alcohol & Tobacco	0.612					
Clothing	0.875					
Housing	1.236					
Services	1.246					
Others	1.877					

Data comes from the Spanish Permanent Survey of Consumption (EPC), years 1977 to 1983, and we select families reporting full information for at least 12 consecutive quarters (3 years). The estimates also include the age of the family head and its square, the log of family members, a time trend and dummy variables for the quarter ref.: quarter 4

zero for our estimated figures; that is to say, the compensated own-price elasticity in our sample of families is, on average, significantly equal to the income elasticity.

We then calculate the condition using the compensated cross-price elasticities for all the goods of the system by using the following σ_k for food, alcohol and tobacco, clothing, housing, services and other goods, 0, 1, 0.75, 1, 0.75 and 0.75, respectively. We prove that condition (4) holds in only 585 observations using the parameter estimated with the QUAIDS model. These families are characterized by having per capita income in the two highest centiles of the distribution, showing no differences in family size. Those households have a higher per capita expenditure, indicating a general negative relationship between per capita food expenditure and family size for rich families. Those rich households are also characterized by a higher proportion of household heads with tertiary education, and are less likely to reside in rural areas.

Conclusions

We explore the relationship between family size and food consumption by using a Spanish panel data containing rich information about consumption. Deaton and Paxson (1998) used cross-sectional data to examine expenditure data from a range of developed (US, UK, France) and developing countries (Thailand, Pakistan, South Africa) and found that, since there are few substitutes for food, the price elasticity of food is low-, and the-income effect dominates.

We particularly take advantage of panel data to estimate a flexible demand system, QUAIDS, from which we derive the price and income elasticities, which allow us to test the theoretical condition derived from the existence of economies of scale. Results indicate that this condition expresses a positive relationship between per capita food expenditure and family size holds in only 585 observations, with these families being characterized by having per capita income in the two highest centiles of the distribution. The economic behavior of families depends on the specific economic situation of the country and, consequently, our results should be interpreted in the context of the particular economic crisis of the 1970s and 1980s, as a consequence of the economic shocks due to the very high oil prices that affected Spain and the rest of world.

Data Availability Available upon request.

Code Availability Available upon request.

Declarations

Conflict of interest Authors declare that they have no conflict of interests.

References

Abdulai, A. (2003). Economies of scale and the demand for food in Switzerland: Parametric and nonparametric analysis. *Journal of Agricultural Economics, 54*, 247–267.

Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics, 79*, 527–539.

Barten, A. P. (1964). Family composition, prices and expenditure patterns. In E. G. Mills, J. K. Whitaker, & P. Hart (Eds.), *Econometric analysis for national economic planning.* Springer.

Brzozowski, M., Crossley, T. F., & Winter, J. K. (2016). Does survey recall error explain the Deaton-Paxson puzzle? *Economics Letters, 158*, 18–20.

Crossley, T., & Lu, Y. (2018). Returns to scale in food preparation and the Deaton Paxson puzzle. *Review of Economics of the Household, 16*, 5–19.

Deaton, A., & Paxson, C. (1998). Economies of scale, household size and the demand for food. *Journal of Political Economy, 106*, 897–930.

Deaton, A., & Paxson, C. (2003). Engel's what? A response to Gan and Vernon. *Journal of Political Economy, 111*, 1378–1381.

Gan, L., & Vernon, V. (2003). Economies of scale, household size and the demand for food. *Journal of Political Economy, 111*, 1361–1377.

Gibson, J. (2002). Why does the Engel method work? Food demand, economies of size and household survey methods. *Oxford Bulletin of Economics and Statistics, 64*, 341–359.

Gibson, J., & Kim, B. (2007). Measurement error in recall surveys and the relationship between household size and food demand. *American Journal of Agricultural Economics, 89*, 473–489.

Gibson, J., & Kim, B. (2018). Economies of scale, bulk discounts, and liquidity constraints: Comparing unit value and transaction level evidence in a poor country. *Review of Economics of the Household, 16*, 21–39.

Griffith, R., Leibtag, E., Leicester, A., & Nevo, A. (2009). Consumer shopping behavior: How much do consumers save? *Journal of Economic Perspectives, 23*, 99–120.

Horowitz, A. (2002). Household size and the demand for food: A puzzle resolved?. Paper presented at the Northeast Universities Development Consortium Conference, Williams College.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.