Extended Families across Mexico and the United States

Extended Abstract

PAA 2013

Gabriela Farfán

Duke University

After years of research we've come to learn quite a lot about household allocation decisions. For a long time households were modeled as if they were individuals, abstracting from the underlying decision problem that involves a number of individuals with presumably different preferences. This traditional model, known as the Unitary Model of the household, would predict, for instance, that all that matters is total resources but not the intra-household distribution of resources. However, there is now substantial evidence rejecting such assumption. Household-level outcomes are the result of a complex decision process, where the individual distribution of resources, which is thought to correlate with decision power, matters. The natural question that followed was to see whether despite the complexity of this process household members were able to cooperate. For the most part, the literature has failed to reject that household allocation decisions are Pareto efficient.

Also widely recognized in the literature is the fact that households are not isolated units. In particular, extended families play an important role in shaping individual and household decisions. The role of family interactions is particularly relevant in developing settings. In the absence of well-functioning financial and insurance markets, families are found to facilitate investments and engage in informal insurance agreements, among other services[1].

While we know inter-household interactions are important, we know relatively less about how families share and allocate their resources. The lessons derived from the intra-household literature do not necessarily hold. For instance, as households are endogenously formed, i.e. individuals choose who they live with, it is not terribly surprising they have not been found to be inefficient. But the situation is quite different when we look at families. In this case we are thinking about interactions between family members that do not live together, and therefore share less information and interact less frequently. Furthermore, if we think about family

[1] See, for example, Rosenzweig and Stark 1989, Angelucci et al. 2010, Thomas and Frankenberg 2007.

members living far away from each other, we can imagine asymmetric information could have an important role to play in preventing families from reaching efficiency.[2]

The objective of this project is to look at how families, with members split across different households, share and allocate resources. We start by analyzing whether resources of non-co-resident family members matter. Conditional on being relevant, we test whether families share resources completely, or are altruistic. This would be the Unitary model of the family, where family resources have the same effect on outcomes regardless of whether they come from within or outside the household. Finally, we apply the Collective model to the family decision problem, and explore whether the allocation of resources among families is consistent with Pareto efficiency.

Given the informal nature of family contracts, the availability of commitment and monitoring devices will certainly determine the scope for cooperation, both of which are influenced by the information available to family members. Therefore, one could imagine that if there is a situation where Pareto efficiency was hard to achieve, it would be in a migration setting. In this scenario, interactions among family members are infrequent and barriers to information are greatest. We also explore this hypothesis by exploiting variation in the geographic dispersion among family members, the timing in migration histories, and the degree of observability in different outcomes.

We look at these questions in the context of Mexico, a developing setting where both internal and international migration are prevalent. Migration to the US is a central element to the Mexican economy. Estimates suggest there are about 12 million Mexican-born individuals in the US, who account for about 10% of the Mexican population (Pew Hispanic Center 2009). It is estimated that, in 2010, remittances worth US$22 billion were sent from the U.S. back to Mexico, which places Mexico as the third largest recipient of remittance income across the globe, behind China and India (World Bank 2011). Characteristic of the Mexico-US setting are the high rates of circular and temporal migration, as well as the existence of important migration networks operating in the host country. All of these elements illustrate a setting where an important share of Mexican households has relatives living in the U.S., migrants retain close links with the home country, and there is an important flow of information across the border.

For these reasons, it is of great importance for this project that the data we use has information on family members living in Mexico as well as on family members living in the US. The data we use is the Mexican Family Life Survey (MxFLS), an ongoing longitudinal survey that collects extensive information on individuals, households, families and communities. The first wave was conducted in 2002, the second wave implemented in 2005-2006, and the third one is in the final stages of field-work. At baseline, the sample consists of 8,440 households spread

[2] Related to the consequences linked to the migrant's lack of control over the use of remittances see Chen 2006, de Laat 2005, Ashraf et al. 2011

out across 150 Mexican communities, and is representative at the national, rural-urban and regional levels.

Central to this study is the follow-up policy of the survey, as it allows us to link individuals in different households who belong to the same family, and it provides us with the geographic variation we need to study migrant families. By design, every individual interviewed at baseline, as well as any child of these individuals born after 2002, is sought for interview in every follow-up. Following Altonji et al. 1992, members who split from their original households are linked to their root or baseline household, to define a family or dynasty. Additionally, individuals or households that move are interviewed in their new location, either within Mexico or in the United States. The interview of respondents in the U.S. is a distinctive and unique feature of these data. Many studies collect information on international migrants by asking other household members, but few large-scale household surveys have tried to follow migrants across international borders. The U.S. component of the survey includes a very comprehensive set of modules that follow closely those applied in Mexico, at the same time that they incorporate specific changes to capture the particularities of the life of Mexicans in the United States. As a result, these data provides us with a sample of families, some of which have members interviewed in different households and different locations, including members interviewed in the U.S.

Taking advantage of this rich data, we classify families in different types based on the geographic dispersion among their members, and see whether different patterns emerge when we move from families where all members live in the same locality (“*neighbor” families*), to the extreme where the family has members located in Mexico and in the U.S (*international families*), going through cases in between.

We start by testing the role of family resources on two sets of outcomes: household budget shares and child human capital. Both of these outcomes have been extensively used in the literature, and we think it is important to look at them simultaneously, as the conclusions we reach might not be the same across all of them. For instance, one might imagine family members not caring too much about how much another individual spends on clothing, but it might be very relevant if children go to school. In a similar way, it might be harder for family members to monitor some outcomes than others.

1. Model

Here we present a simple representation of the collective model developed in the intra-household literature applied to the family decision problem (Chiappori and coauthors). This will allow us testing whether the family allocation of resources is consistent with Pareto efficiency.

Let W represent family welfare, a function of the utility U^h of each household h, with h=1, …,H, H the size of the family. Let $q \in R^I$ be the vector of consumption goods, q_i^h denote

consumption of good *i* by household *h*, $Y = \sum_{h=1}^{H} y_h$ denote total resources, and *a* and ϵ denote vectors of observable and unobservable preference factors. If resources are shared efficiently within the family, household demands are the solution to the following problem[5]:

$$(1) \qquad \max_{q^1, \dots, q^H} \quad \mu^1 U^1(q^1; a, \epsilon) + \cdots + \mu^H U^H(q^H; a, \epsilon)$$

$$s.t. \;\; p \sum_h q^h \leq \sum_h y_h = Y$$

where $\mu_h(p, y, z, a, \epsilon)$ represents the Pareto weight attached to household *h*, with $\sum_{h=1}^{H} \mu_h = 1$, $y = (y_1, \dots, y_H)$, p is a vector of market prices, and U^h represents well-behaved "household preferences"[6]. The vector z denotes distribution factors, meaning variables that do not affect preferences nor they affect the budget constraint but modify household demands through their effect on the distribution of power within the family.

Problem (1) can alternatively be solved in two stages[7]. In the first stage, the family agrees on a sharing rule θ that assigns to each household a share of total resources, the distribution of resources being a function of μ. In the second stage, each household solves the following problem:

$$(2) \qquad \max_{q^h} \quad U^h(q^h; a, \epsilon) \qquad s.t. \;\; p\, q^h \leq y_h^* = \theta_h(p, y, z, a, \epsilon)\, Y$$

$$\theta_h \in (0,1) \qquad \sum_h \theta_h = 1$$

Denote the solution to these individual problems with $q^{h*} = g(p, y_h^*; a, \epsilon)$. From this expression it is clear to see that:

$$(3) \qquad \frac{\partial q_i^{h*}/\partial y_k}{\partial q_i^{h*}/\partial y_l} = \frac{\partial q_i^{h*}/\partial y_h^* \;.\; \partial y_h^*/\partial y_k}{\partial q_i^{h*}/y_h^* \;.\; \partial y_h^*/\partial y_l} = \frac{\partial q_j^{h*}/\partial y_k}{\partial q_j^{h*}/\partial y_l}$$

for any two households k, l in family f. That is, while own household resources and family resources are allowed to have different effects on household demands, they satisfy a very particular restriction: the ratio of marginal effects of any two sources of income is the same across all goods (the ratio is independent of good *i*). In the empirical section, we use these conditions to test for Pareto efficiency within families.

2. Data

[5] At this point, the model is a static model. Extensions to include dynamics are left for future work.

[6] This would be the case if household preferences can be written as a weighted sum of individual preferences with *fixed* weights (consensus model *a la* Samuelson 1956), every household member has identical preferences, or the weights of all but one household member are set to zero (dictator model).

[7] See Chiappori 1992 and Bourguignon et al. 2009.

The data used in this project is the Mexican Family Life Survey (MxFLS), an ongoing longitudinal survey that collects a rich set of information on individuals, households, families and communities. The first wave, conducted in 2002, includes 35,677 individuals in 8,440 households spread out across 150 Mexican communities. At baseline, the sample is representative at the national, rural-urban and regional level. The second wave of the survey was implemented in 2005-2005, reaching a 90% overall re-contact rates. The third wave is in the final stages of field work, and by now we have an 85% re-contact rate.

As briefly mentioned in the introduction, we use the panel structure of the survey to identify extended families in our data. By design, MxFLS tracks every member interviewed at baseline in 2002, as well as every child of original household members who are born after 2002. From now on we will refer to any individual interviewed at baseline, or child of such individuals born after 2002, as panel members, and the 2002 household as original or root household. In later rounds, if any panel member is not part of the original household at the time of the follow-up, that individual, together with her/his new household members are interviewed as a new household. In this way, following Altonji et al. 1992, we link every split-off household in our data to their root household, and define this group as a family or dynasty.

A distinctive feature of the data is the fact that panel members living in the U.S. at the time of the follow-up are also followed and interviewed in their new household. Many studies collect information on international migrants from other household members, but few large-scale surveys have tried to follow migrants across international borders[10]. In the second wave MxFLS interviewed, mostly by phone, 91% of those believed to be in the US at the time. In the third wave, we interviewed face to face 85% of the panel migrants living in the US. Important for our analysis, the US component of the survey includes a very comprehensive set of modules that follow closely those applied in Mexico, at the same time that they incorporate specific changes that capture the particularities of the life Mexicans have in the United States.

Considering the way families are identified in the analysis, it is important to look at the definition of *household* adopted in the survey. In particular, the definition of household at baseline and the definition of household in the U.S. are relevant.

The first definition is important because we will be analyzing the interaction between root and split-off households. If original household members do not keep any link once they split, we would not be looking at a relevant unit. The definition of household applied in Mexico is that of "living together and eating from a common pot". For every split-off in the third round, we checked the relationship that this individual had in 2002 with the household head, and found

[10] Virtually all Mexican migrants remain in Mexico or go to the United States. In our data, we only have 18 individuals who migrated outside of these two destinations (for example, to Canada, or the United Kingdom).

that virtually everyone is a close relative, with children and grand-children accounting for the great majority (almost 80%)[11].

The second definition that deserves explanation is that of household in the U.S., which was a challenging component when designing the U.S. component of the survey. The reason is that it is not uncommon to find migrants living together with the only purpose of saving on rent (and utilities), but who otherwise have nothing in common. As a result, had we followed the traditional definition to identify household members we would have information on a unit that is of little interest for economic analyses. Thus, we added an additional condition to the usual definition of a household. As stated in the questionnaire, a household is "a group of individuals who usually live together, usually consume meals provided by a common budget *and* usually share *other* expenses" (besides housing and food). The data show that almost 60% of the households share their dwelling with non-household members as defined above. Looking at Table 1, we see how the average number of individuals per dwelling is almost 5, while the average household size is only 2.7. The bottom panel of the table also shows in detail the relationship of household members to the household head. With the stated definition of a household, the resulting household structure seems to be quite standard. Over 85% of household members are spouses and children, and the other 25% corresponds mostly to parents, siblings, grandchildren, nephew/nieces.

4.1 Sample of families

In Table 2 we present some statistics that illustrate the basic structure of the data. In the third wave we have at this point 9,813 households, 731 of which were interviewed in the Unites States, and roughly half of which belong to families with at least two households in the data. Among the 1,937 families with more than one household, 27% have at least one household in the US.

In order to explore the importance of information asymmetries due to migration effects, we are going to stratify our sample based on the geographic location of family members, and see how the results vary with that. We will present the results for five groups: all families, only families who live in the same locality ("*neighbor" families*), families who have members across different localities, families who have members across different states, and finally families who have members both in Mexico and in the U.S (*international families*).

4.3 Outcomes of interest

As mentioned in the introduction, we estimate the model on two sets of outcomes: household budget shares and child outcomes.

[11] "Worker" or "other" account for 1%.

Having expenditure data on households located in Mexico as well as on their relatives located in the U.S. is a very unique feature of our data. We designed the consumption module of the U.S. questionnaire following the one applied in Mexico, which allows us to define the same bundles of goods and construct comparable expenditure shares across all households in our sample[13].

We divide total expenditures in 3 groups: food (includes both food consumed at home and meals outside); non-food (includes personal care, clothing, health, education, recreation, house cleaning, semi-durables and other, communication and transportation); and housing (which includes rental value and utilities)[15]. We are trying with more disaggregated groups, taking into account the nature of the goods and the share they represent on total expenditures, but this classification seems good enough to convey the main message of the analysis. We convert all magnitudes to monthly expenditures, and units are measured in dollars, using PPP exchange rates when reported in Mexican pesos. Table 4 presents some summary statistics for all families, as well as by type of family.

With respect to child outcomes, we start with three markers of human capital. Two relate to the nutritional status of children, Height-for-age and BMI-for-age zscores, and the third one is years of education. Both the Mexican and the U.S. components of the survey include a health section, which includes several health markers for all household members. All measures are taken by trained personnel. Table 4 presents summary statistics on children 0 to 9 for the two nutritional outcomes, and children 6-16 for years of education.

3. Empirical Implementation

From the model presented above we derived the following conditional demand functions:

$$q^{h*} = g(p, y_h^*(p, y, z, a, \epsilon, Y); a, \epsilon), \text{ with } y = (y_1, \dots, y_H) \tag{11}$$

In a first attempt to test the model we estimate the following linearized version of demand functions:

$$out_{hf}^i = \alpha^i + \beta_1^i \log(y_h) + \beta_2^i \log(y_e) + x_{hf}'\gamma^i + \varepsilon_{hf}^i \ , \tag{12}$$

[13] The modules across the two countries are not exactly the same. Two main differences are the level of disaggregation in the questions (e.g. all food versus types of food), and for a few items (education and semi-durables) the reference period, which in the US does not go beyond 3 months prior to the interview date. In terms of the level of disaggregation, when designing the questionnaire we faced the trade-off between comparability with the module in Mexico and the length of the interview in the US. Since it was the first experiment with a face-to-face interview, and the questionnaire is quite long, we compromised selecting broader categories. With respect to the reference period, we did not want to ask for a period too long before the interview date in order to minimize the probability that the reference period covers both time spend in the U.S. and time spent in Mexico for very recent or circular migrants . Even though we expect these two differences to affect reported expenditure, the assumption we need is that expenditure shares are not affected.

[15] In the case of home owners, we use the self-reported rental value of their dwellings.

where out_{hf}^{i} is outcome i of household h in family f, y_h are household resources, y_e are extended family resources (*total family resources less household resources*), and x'_{hf} is a vector of household and family characteristics. Resources will be measured by the log of per-capita expenditures[16].

After estimating system (12), the tests to be performed are:

Unitary test: $\beta_1^i = \beta_2^i$

Pareto test: $\beta_1^i/\beta_2^i = \beta_1^j/\beta_2^j$ for any pair of outcomes i, j.

We estimate a seemingly unrelated regressions demand system (SUR), with cluster standard errors at the family level. To test Pareto efficiency we need to implement cross-equation tests. We estimate non-linear Wald tests calculated using the delta method allowing for clustering at the family level. We re-express the test as the cross-product of coefficients instead of ratio of coefficients. We present the results of pair-wise comparisons across any two outcomes as well as the joint test for all pairs simultaneously.

4. References

Altonji A., F. Hayashi and L. Kotlikoff (1992). "Is the Extended Family Altruistically Liked? Direct Tests Using Micro Data", *The American Economic Review*, 82(5) , pp. 1177-1198

Angelucci M., G. De Giorgi, M. Rangel and I. Rasul (2010). "Family networks and school enrolment: Evidence from a randomized social experiment", *Journal of Public Economics*, 94, pp. 197-221

Ashraf N., D. Aycinema, C. Martinez and D. Yang (2011). "Remittances and the problema of control: A field experiment among migrants from El Salvador", Universidad de Chile, Mimeo

Bourguignon F., M. Browning and PA Chiappori (2009). "Efficient Intra-Household Allocations and Distribution Factors: Implications and Identification", *Review of Economic Studies*, 76, pp. 503-528

Chen J. (2006). "Migration and Imperfect Monitoring: Implications for Intra-Household Allocation", The American Economic Review Papers and Proceedings, 96(2), pp.227-231

Chiappori PA (1992). "Collective Labor Supply and Welfare", *Journal of Political Economy*, 100(3), pp. 437-467

De Laat J. (2005). "Moral Hazard and Costly Monitoring: The Case of Split Migrants in Kenya", Job Market Paper

[16] Throughout the paper, *family resources* refers to the sum of household resources for the complete family, while *extended-family resources* are total family resources minus own household resources.

Rosenzweig M. and O. Stark (1989). "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India", *Journal of Political Economy*, 97(4), pp. 905-926

Thomas D. and E. Frankengerg (2007). "Household Responses to the Financial Crisis in Indonesia: Longitudinal Evidence on Poverty, Resources and Well-Being", in Ann Harrison, ed. *Globalization and Poverty*, University of Chicago Press

Thomas D. (1990). "Intra-Household Resource Allocation: An Inferential Approach", *The Journal of Human Resources*, 24(4), pp 635-664

Table1: Living arrangements in the US

Dwelling characteristics		
	mean	**sd**
# individuals in dwelling	4.94	2.7
# relatives in dwelling	3.77	2.45
household size	2.71	1.64
Total households	599	
Relation to head of household		
	Freq.	**Percent**
Head	596	36.95
Spouse	**312**	**19.34**
Son/Daughter	**551**	**34.16**
Step child	**35**	**2.17**
Son/Daugher in law	9	0.56
Father/Mother	8	0.5
Father/Mother in law	7	0.43
Brother/Sister	**31**	**1.92**
Brother/Sister in law	14	0.87
Grandson/daughter	**15**	**0.93**
Uncle/Aunt	1	0.06
Nephew/Niece	**19**	**1.18**
Cousin	6	0.37
Not relative	5	0.31
Other	4	0.25
Total individuals	1,613	

Table 2: Basic Structure of the Data -MxFLS3

HOUSEHOLDS		FAMILIES	
# Households		**# Families (Dynasties)**	
In MX	9,113		
In US	739		
Total	**9,852**	***Total***	**7,136**
# HHS with extended family		**# Families w/at least 2 households**	
In MX	3,977	*All hhs in Mx*	1,415
In US	688	*At least one hh in US*	534
Total	**4,665**	***Total***	**1,949**

Table 4: Summary statistic of main variables, by family type

	All families			Different locality			Different State			International families			Neighbor families		
	mean	sd	# obs	mean	sd	# obs	mean	sd	# obs	mean	sd	# obs	mean	sd	# obs
Sample of households															
food share	49.86	15.99	4408	47.54	16.63	2074	46.00	16.89	1348	45.16	16.73	1242	51.91	15.11	2334
non-food share	26.06	14.92	4408	27.93	15.26	2074	29.29	15.41	1348	29.78	15.42	1242	24.40	14.41	2334
housing share	24.06	12.94	4408	24.48	13.29	2074	24.64	13.67	1348	24.98	13.73	1242	23.69	12.61	2334
log household pce	5.48	0.86	4407	5.65	0.97	2073	5.79	1.04	1347	5.85	1.04	1241	5.34	0.73	2334
log family pce	5.50	0.63	4128	5.70	0.61	1880	5.87	0.55	1188	5.94	0.51	1087	5.33	0.60	2248
log extended-family pce	5.49	0.79	4127	5.69	0.85	1880	5.85	0.88	1188	5.93	0.86	1086	5.33	0.69	2247
Sample of children															
height-for-age	-0.39	1.20	3646	-0.29	1.25	1549	-0.32	1.22	968	-0.31	1.22	864	-0.46	1.16	2097
bmi-for-age	0.34	1.33	3623	0.42	1.40	1542	0.41	1.40	962	0.45	1.40	860	0.27	1.27	2081
years of education	4.45	3.10	2887	4.64	3.11	1318	4.82	3.09	882	4.80	3.09	799	4.29	3.08	1569
log household pce	5.11	0.76	5904	5.17	0.83	2637	5.20	0.89	1690	5.26	0.90	1516	5.06	0.70	3267
log family pce	5.31	0.61	5451	5.52	0.59	2330	5.69	0.50	1425	5.78	0.45	1261	5.15	0.58	3121
log extended-family pce	5.45	0.79	5449	5.72	0.85	2330	5.98	0.82	1425	6.09	0.77	1259	5.24	0.67	3119