SUPPLEMENT TO "ECONOMICS AND MEASUREMENT: NEW MEASURES TO MODEL DECISION MAKING"

A1 Supporting References

This appendix section provides details on the relevant literature that was not presented for the different topics discussed in Sections 2 and 3 of the main text. This was not included in the main text because of page length limitations. We report the supporting references organized along the topics from Sections 2 and 3.

A1.1 The Context

In addition to the studies cited in the main text, we suggest the following references:

Contingent valuation and hypothetical choices. A good survey of contingent valuation's use in other disciplines such as marketing is in Louviere et al. (2000) and Carson (2012). Additionally, several studies, including List and Gallet (2001) and Murphy et al. (2005), discuss common biases in answers to hypothetical situations questions.

Stated preferences, conjoint analysis and choice data. An example of the arguments about what and how to measure can be found in Luce (1959), Luce and Tukey (1964), and Luce and Suppes (1965). The reliability and predictive power of buying intentions and purchasing probabilities data are discussed in Juster (1966). Curtin (2016) provides a nice survey. Studies of stated preferences and answers to hypothetical questions include Blass et al. (2010), Kesternich et al. (2013), and Ben-Akiva et al. (2019). A recent example which considers explicitly how data on hypothetical choices can be used to identify causal links in economic models is Bernheim et al. (2021). Other related studies are Harris and Keane (1998), Erdem et al. (2005).

Experimental economics. Evidence on measures of preference and attitudes towards redistribution, attitudes towards migrants, bargaining and social preferences, reciprocity in conflict areas, and willingness to compete through experiments combined with observational data are reported in Almås et al. (2020b), Alesina et al. (2018), Almås et al. (2018), Buser et al. (2014), Cavatorta and Groom (2020). There is also evidence on how actual data have been reproduced in the lab, in particular for the analysis of auction models, for instance, Ertaç et al. (2011); Salz and Vespa (2020).

Subjective expectations. An example of early collection of subjective expectations is Visco (1984). Important references include Dominitz and Manski (1997), Dominitz and Manski (1996), Potter et al. (2017). A non-exahustive list of studies using subjective expectations data include Jappelli and Pistaferri (2000), Pistaferri (2001), Pistaferri (2003), Van der Klaauw and Wolpin (2008), Kaufmann and Pistaferri (2009), Attanasio and Augsburg (2016), Paiella and Pistaferri (2017), Attanasio et al. (2018), and Giustinelli et al. (2019). A recent example of high quality subjective expectations data are those collected by the Federal Reserve Bank of New York: the *Survey of Consumer Expectations* (https://www.newyorkfed.org/microeconomics/sce/background.html).

Chen et al. (2020) show that, without assuming rational expectations (or data on subjective expectations), only set (rather than point) identification can be achieved.

Beliefs. In addition to those cited in the text, several studies look at the elicitation of subjective beliefs. For instance, Boneva and Rauh (2018), Dominitz and Manski (1996), Wiswall and Zafar (2015), and Delavande and Zafar (2019) have studied beliefs about the returns to college and college enrollment choices, Bobba and Frisancho (2020) assessed how college application choices are affected by students' perceptions of their own ability and how these can be changed by additional information, while Dizon-Ross (2019) studied how parental beliefs about their children's abilities affect their choices.

A1.2 Measurement and Theory

In addition to the studies cited in the main text, the following references are important for the relationship between theory and measurement:

National accounts. Classic studies that led to the development of Modern National Accounts include Keynes (1936), Kuznets et al. (1937), Kuznets (1941), Gilbert et al. (1949), and Stone (1984). Recent examples for the effects of new products and quality on the measurement of inflation include Bils and Klenow (2001), Bils (2009), and Crawford and Neary (2021); Aghion et al. (2019) for the importance of *creative destruction* to measure growth; and Neary (2004) and Almås (2012) for international comparisons.

Prices. In many contexts, researchers use measures of unobserved latent factors that are the result of complex algorithms and procedures that summarise very large raw data. An obvious example is that of consumer price indices, whose weights to aggregate a myriad of different prices are based on certain models of demand behavior and are periodically updated using new data on consumption baskets purchased by consumers. The formulation of the '*measurement system*', mapping available data (the prices of many individual commodities) into a construct of interest (a price index) is driven by specific theoretical models of individual behavior, and is periodically updated to reflect changes in behavior caused, for example, by the introduction of new commodities or improvements in their quality. For price indices, see Stone (1954), Christensen et al. (1975), Deaton and Muellbauer (1980), and more recently Nordhaus (1998). However, in some contexts prices may not be linear and change with the quantity purchased, as in many models of price discrimination (e.g., Maskin and Riley (1984), Jullien (2000), and Attanasio and Pastorino (2020)). Recent progress has been made with the analysis of scanner data (Einav et al. (2008), Griffith and O'Connell (2009), Dubois et al. (2020)).

Multiple Indicators Multiple Causes (MIMIC) models. References on this include Wright (1934), Duncan (1966), Goldberger (1971), Goldberger (1972), Griliches (1974), Jöreskog and Goldberger (1975), and Chamberlain and Griliches (1975).

Labor markets and production functions. Studies of models to explain recent increases in inequality complemented by innovative data include, for instance, Deming (2017), Acemoglu and Restrepo (2019), Bloom et al. (2019), and Scur et al. (2021).

Decision units. A number of important theoretical results have been derived within the collective model of Chiappori (1988, 1992); Browning and Chiappori (1998); Bourguignon et al. (2009); Cherchye et al. (2011). Attanasio and Lechene (2014) present tests of the collective model using the variability induced in a standard household demand curve by two different *distribution factors*, i.e., variables that affect Pareto weights but not utility. Characterizing the parameters of these models and testing their validity exclusively with *choice data* on household consumption and expenditure data might be challenging. A number of papers use individual-level data within households,

for instance Dunbar et al. (2013); and more recently Lechene et al. (2022), though strong assumptions are required to identify the determinants of individual behavior.

Elicitation of preferences and beliefs. With an *experimental approach*, researchers can place individuals in identical settings – and hence should have the same beliefs – and then elicit preferences through real (incentivized) choices. One example of such an approach in the literature on preferences for redistribution is Almås et al. (2020b). A similar approach has also recently been used in the study of household decision making: Cherchye et al. (2021) studies actual choices in an economic experiment for married couples in Nairobi, Kenya. In the experiment, parents were asked to allocate money between themselves and nutritious meals for one of their children. From these choices, the authors test models of household decision making and they estimate a parametric (structural) model of household collective decision making. Through this, they reveal preferences and Pareto weights related to bargaining power. The stated preference method instead fixes beliefs by giving respondents a set of hypothetical scenarios or choices where there is full information about the relevant parameters in each scenario. An early example of such a strategy can be found in the aforementioned Juster (1964) and Juster and Shay (1964). More recently, Ameriks et al. (2020) and Bernheim et al. (2021) have used hypothetical questions to estimate parameters that characterize individual preferences. Within such an approach, one can use answers to hypothetical questions to identify *preferences*, so that beliefs can then be inferred when these data are considered jointly with *choice data*. Furthermore, instead of eliciting beliefs (preferences) and inferring preferences (beliefs) as a residual by combining elicitation choice data, one may directly elicit *both* beliefs and preferences in particular choice situations, where both heterogeneities in tastes and beliefs might be important. Adams and Andrew (2019) suggest using *survey experiments* to elicit *average* beliefs and preferences in the Indian (specifically Rajasthani) marriage market for young brides. In addition to the papers cited in the main text, List et al. (2021) have elicited beliefs from parents in Chicago. Bobba and Frisancho (2020) collect and use data on self-perceptions of academic achievement among high school students in Mexico; Miller et al. (2020) study beliefs about contraception effectiveness.

The environment and social capital. Evidence on studies that have looked at different approaches to measure social capital, including measures of participation in certain activities (from blood donation to church attendance, see, for example, Guiso et al. (2004) and Guiso et al. (2006)) to data derived from field experiments (see Attanasio et al., 2012) on determinants of group formation, to the effect of deterrence on preferences (see Cavatorta and Groom, 2020).

Measures of child development. The UN has facilitated the development of the Multiple Indicator Cluster Survey (MICS), which has recently been discussed in Bornstein et al. (2021). The Gates Foundation has funded a large effort to pull together a number of new indicators that could be comparable across contexts and countries and would be relatively easy to administer. This has given rise to the Global Scale for Early Development (GSED) initiative, discussed in GSED (2021) and Black et al. (2019). McCoy et al. (2021) describe the construction and use of the CREDI questionnaire.

New measures. Examples of new measures include Ferrario and Stantcheva (2022) on using text analysis of open-ended questions and the work on perceptions about several issues (Alesina et al., 2023; Stantcheva, 2022; Alesina et al., 2021).

A2 Specification of Measurement Systems

This Appendix provides more details on the general issues in the specification and estimation of measurement systems that were discussed in Section 5.1 of the main text.

One assumption that can be relaxed in the context of the measurement system, and that we mention in the main text, is that of linearity, especially when the available measures are binary or discrete. Here, we explain the logic. In such a situation, a common approach model is the Item Response Theory (IRT). In the case of binary measures, an IRT model relates a factor to a latent measure, which in turn determines the observed outcome. Omitting the superscript for the latent factor θ, we have:

$$m_{it}^k = \begin{cases} 1 & if \;\; \alpha_t^k + \beta_t^k \theta_{it} + \varepsilon_{it}^k > 0 \\ 0 & otherwise \end{cases} \tag{A1}$$

The specific IRT model is then determined by assumptions about the distribution of the measurement error term ε_{it}^k; assuming normality, one gets a *Probit* type of model, while assuming a logistic distribution one obtains a logistic relation, or what is usually referred to as a Rasch model. What we are considering is often defined as a 2-parameter Rasch model. Restricting β_t^k to 1, one obtains a 1-parameter Rasch model. In the literature, a 3-parameter Rasch model considers the possibility of random 'correct' answers through an additional parameter.

Another more recent example of IRT use, whose use has become more common in economics, is that of polygenic scores, which aggregate data from many different sites of the human genome and are based on correlations from a wide population data set. Lee et al. (2018), for instance, present estimates of a polygenic score which is associated with individual educational attainment in a specific population. Interestingly, several authors recently noticed that several estimates of the same polygenic score might be available, where the weights to aggregate information from individual loci are based on different samples and/or slightly different methodologies. These alternative estimates can then potentially be used to deal with measurement error problems, as in the model discussed above. Relevant contributions on the literature about the identification of measurement systems models are Schennach (2004) and Hu and Schennach (2008).

A3 The Tanzania data, new measures and results

The study "Improving children's life chances in high-risk, low-income settings: Designing a "new generation" longitudinal cohort study of child development", which was approved by the University College London Ethics Committee (2168/013) and the National Institute for Medical Research and the Ministry of Health, Community Development, Gender, Elderly Children of Tanzania (NIMR/HQ/R.8c/Vol. I/1254), was conducted in several stages (see Figure A1) (see the Tanzania Field Pilot Report Augs-

burg et al. (2019) for more information). The study sample was randomly chosen to be representative of the total population. The villages were selected from a random sample in the Kagera region. In the selected villages, all households were listed, and from those satisfying the selection criteria, a random sample was drawn.

The Kagera region is located in the northwest corner of Tanzania on the western shore of Lake Victoria. The population is around 3 million people which is about 5 percent of the national population. It has an intercensal growth rate 2002/2012 of 3.2 (2.7 average for Tanzania) and the double percentage of children age 6-59 months classified as having malaria, according to the Rapid Diagnostic Test in 2017 in terms of country percentage (15.4% vs 7.3%). The Kagera region is predominantly rural, and agricultural-dependent, producing mainly bananas and coffee in the North and rain-fed annual crops in the South.

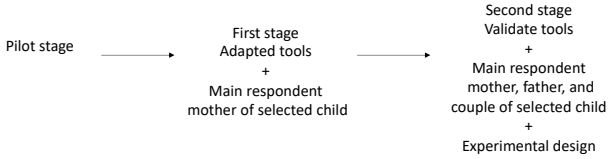
During the pilot period a number of well-known and tested measures, were adapted to the local context using multiple techniques (focus group discussions, ethnography study, and pre-pilot in small samples). We used a multi-method approach for the early childhood development (ECD) instruments, combining three main paths: parent/observer report, direct assessment, and observation (through filming and coding).

After the pilot period, the first stage was implemented in June 2018. Approximately 450 households were recruited across 10 villages. The sample households were randomly selected among those with a child aged between 6 and 36 months and a mother aged between 15 and 25. In this sample, child development was measured using several tools. In addition to standard demographic information about the households involved, the survey also included measures of parental investment.

Information collected during a partial census was used to randomly sample pairs of primary caregivers and children of eligible households from each of the following groups: 150 primary caregivers aged 15-25 years with a child aged 6-11 months, 150 primary caregivers aged 15-25 years with a child aged 12-23 months, and 150 primary caregivers aged 15-25 years with a child aged 24-36 months.

In a second stage, which was implemented in August 2018, an additional 450 households were recruited according from the same villages the following procedure. First a random sample of 150 households was selected from a list of 5200 households with

Figure A1: Study Design



children aged 6 to 36 months and mothers aged 15 to 25 years in 8 villages. This is what we will refer to as the mothers sample. In addition, a second list of around 2000 households with children aged 6 to 36 months and the fathers present was identified in 5 different villages.[22] From this list, which also included mothers older than 25, an additional sample of households with fathers present and mothers aged between 15 and 25 years was selected and formed what we label as the couples sample. From the remaining households in the second list, a third sample of 150 households was selected that we label as the fathers sample. Because of the way the second list was formed, the fathers sample comprises couples who are considerably older. Indeed, only 6.6% of the mothers in the sample were younger than 25.[23]

The data collected in the second stage is what we mostly use in this appendix and in Sections 5 and 6. The main goal of the second stage data collection was to validate some of the child development measures constructed with the factor analysis of first stage data. However, we also collected a number of new variables aimed at measuring individual tastes, beliefs, and bargaining power within the couple. Each sample's label corresponds to the respondents in this new set of questions. In the mothers sample, for instance, the questions about tastes, subjective beliefs about parental investment returns, and bargaining power were answered by the mother in private; in the couples sample, the taste questions were answered jointly by the couple while the others by the mother in private; in the fathers sample, fathers answered the bargaining power and taste questions in private. All other (standard) modules were answered by the mother.

All data collection followed a rigorous process of tools, instruments, and survey

[22]The two sets of villages belonged to different wards, the Kashai and the Bakoba. Both wards belonged to the Bukoba municipality in the Kagera region.

[23]In the mothers and couples samples, there are a few mothers older than 25.

development.[24] Of the 450 pairs recruited for the last stage of the survey, we obtained usable information from 423 households, comprising 145 in the mothers sample, 136 in fathers sample, and 142 in the couples sample.

The fact that the questions that elicit information on tastes, subjective beliefs about parental investment returns, and bargaining power within the couple are answered by different respondents in the three subsamples makes the survey particularly interesting, as it allows to test the hypothesis that individuals within the family are characterized by differences in these variables. Unfortunately, the recruitment process followed in the field (described above) makes comparisons across different sub-samples difficult to interpret. This is particularly true for the fathers sample, which includes considerably older individuals, as we show below. While the fact that the mothers and couples samples were drawn from different villages is not particularly worrying, given the similarities and the proximity of the villages, the systematic difference in the age structure of the study sample makes comparisons across samples problematic.

In Table A1, we report descriptive statistics on the main features of the three samples discussed above. The table also contains $p - values$ for tests (adjusted for multiple hypothesis testing) of the difference between the mothers and the fathers samples (in column 5) and the mothers and the couples samples (in column 8). In addition to the new measures (discussed below), information on a standard set of variables, including demographics, education, and wealth markers, was also collected in the three samples.

[24] All instruments were translated into the local language, Swahili. Translations were carried out by the field staff following a rigorous back-translation procedure. The household surveys were administered in the child's home by enumerators. The adapted version of the assessments of child development were conducted by trained staff in the presence of the primary caregiver.

Table A1: Household characteristics

	Mothers sample		Couples sample			Fathers sample		
	Mean	Observations	Mean	Observations	Adj. p-value	Mean	Observations	Adj. p-value
Demographics								
Age, M	23.12	145	22.82	142	*0.941*	30.42	136	*0.010*
Age, F	28.98	105	28.83	139	*1.000*	37.05	134	*0.010*
Secondary school or more, M	0.35	144	0.37	142	*1.000*	0.23	136	*0.079*
Secondary school or more, F	0.50	105	0.46	139	*0.990*	0.34	134	*0.059*
Food Share	0.69	145	0.61	142	*0.030*	0.63	136	*0.079*
Wealth index	-0.17	144	-0.04	142	*0.931*	0.22	136	*0.010*
Child characteristics								
Child is male	0.57	145	0.54	142	*0.990*	0.45	136	*0.079*
Age in months	18.44	145	18.25	142	*1.000*	18.36	136	*0.901*
# siblings	0.57	145	0.62	142	*0.990*	2.16	136	*0.010*
Birthrank	2.00	145	2.04	142	*1.000*	3.53	136	*0.010*

Note: The table displays the descriptive statistics for the households."M" indicates values for mothers and "F" indicates values for fathers. Column (5) compares the mothers and couples samples while Column (8) compares the mothers and fathers samples. Columns (5) and (8) reports p-values adjusted for multiple hypothesis testing using the Romano-Wolf approach. Source: Tz Pilot.

Consistent with the sampling scheme, the mothers and couples samples have relatively younger mothers, with an average age of 22, while in the fathers sample, the average mother's age is about 7 years higher. This difference is also reflected in the average father's age in the fathers sample, (between 5 and 7 years higher than in the mothers and couples samples). About 20% of the mothers sample is made of single mothers (i.e., the child's father is not present). The average child's age is uniform across the three samples and the share of male children is lowest in the fathers sample.

The only significant and somewhat surprising difference between the mothers and couples samples is the share of total expenditure spent on food, which could be considered a useful indicator of economic well-being[25] This variable is higher in the mothers sample, at 0.69,indicating a poorer sample, and lowest in the couples sample, at 0.61.

We also observe a number of wealth indicators, which we use to estimate a wealth index, normalized to have zero mean in the whole sample. Consistent with the evidence on the food share, the wealth index is lowest in the mothers sample. These differences in permanent income and wealth, however, are not consistent with the information on education: both mothers and fathers in the fathers sample are the least educated, while the mothers sample is the most educated. These differences might reflect cohort effects.

[25]The intuition for this goes back to the established economic regularity that the food share falls with income (Engel, 1857, 1895). Food shares have been used to identify differences in real income and well-being (Hamilton, 2001; Costa, 2000; Almås, 2012; Almås and Johnsen, 2018).

A3.1 Bargaining Power

In Table A2, we report the average share of the 6600 TSH that the respondents (i.e., the wives in the mothers and couples samples and the husbands in the fathers sample) were willing to forfeit so that the payment would be made to them rather than their spouse.

Using the answers to these questions in the mothers and couples samples (where the wife is the respondent) and the fathers sample (where the husband is the respondent) we estimate a measurement system and, from that, the latent factor of interest, reflecting women's decision power in the couple. The estimates of the measurement system parameters are reported in Table A3.

In particular, in Table A4, we report the results of a regression of the WTP on the latent factor reflecting decision making power, estimated first for the sample of mothers and couples and then for the sample of fathers.

Table A2: Willingness to pay (out of 6600 TSH)

	Wives	Husbands
	Mothers and Couples samples	Fathers sample
Average share	0.320	0.100
Median share	0.061	0.008
Std. Dev.	0.395	0.258
Observations	218	98

Note: The table displays the descriptive statistics for the Willingness To Pay (WTP) variable, which is measured as the share of 6600 TSH that the respondents (i.e., wives in the mothers and couples samples, and husbands in the fathers sample) were willing to forfeit to have the payment made to them rather than their spouse. In the mothers sample, only households where the father is present are included. Source: Tz Pilot.

Table A3: Measurement system for Decision making

	Decision making			
	Wives		Husbands	
	Mothers and Couples sample		Fathers sample	
	β	α	β	α
Own health	1.000	0.000	1.000	0.000
	-	-	-	-
Children's health	1.282	-0.172	1.213	-1.140
	(0.263)	*(0.168)*	*(0.484)*	*(0.249)*
Children's schooling	1.202	-1.254	1.021	-0.680
	(0.380)	*(0.183)*	*(0.419)*	*(0.194)*
Household purchases	0.693	-0.969	1.506	-0.386
	(0.217)	*(0.132)*	*(0.638)*	*(0.242)*
Cooking	0.269	0.906	0.563	-1.039
	(0.116)	*(0.114)*	*(0.296)*	*(0.167)*
Visiting	0.630	-0.933	0.582	0.235
	(0.186)	*(0.127)*	*(0.302)*	*(0.138)*
Factor mean	0.458		-0.217	
Factor variance	0.926		0.402	

Note: This table shows the loading factors (β) and the intercept (α) of an IRT estimated from questions about who is mainly responsible for a number of decisions within the household, including: major household expenditures, children's education, own and children's health expenditures, what food to cook, and whether the wife can go out. Columns (1) and (2) are for the mothers and couples sample, and Columns (3) and (4) for the fathers sample. In the mothers sample, only households where the father is present are included. Standard errors in parentheses. Source: Tz Pilot.

Table A4: Willingness to pay and decision making factors

	Willingness to pay	
	Wives	Husbands
	Mothers and Couples samples	Fathers sample
Decision making factor	-0.081**	0.071
	(0.032)	*(0.054)*
Couple	-0.042	
	(0.054)	
R-squared	0.033	0.018
Observations	218	98

Note: The table displays regressions of the willingness to pay on the decision making variable and a dummy for the couple treatment arm of the experiment. In the mothers sample, only households where the father is present are included. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Tz Pilot.

A3.2 Beliefs on returns to parental investment

In the Tanzania sample, we elicit beliefs about different aspects of the developmental process and the importance of certain parental inputs. In addition to beliefs about cognitive development, we also measure beliefs about the effect of parenting on socio-emotional development, a type of belief that has not been measured before.

Figure A2: Beliefs on Language Development: Returns to Parental Investment

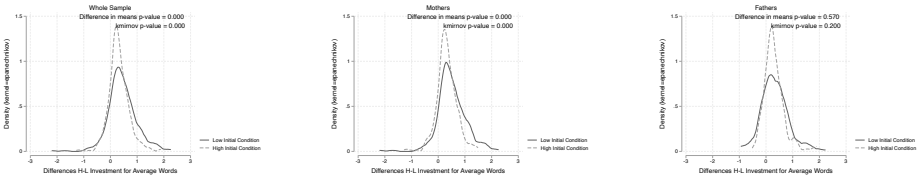
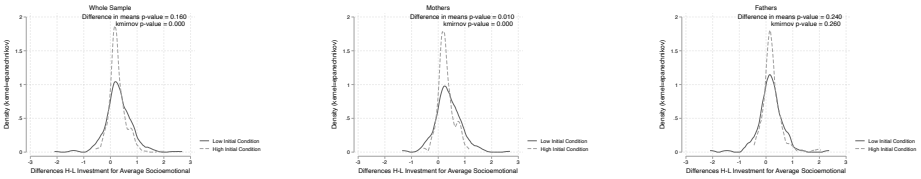


Figure A3: Beliefs on Socio-emotional: Returns to Parental Investment



In Figure A2, we plot the distribution of expected returns under low and high initial conditions for the whole sample as well as for the mothers and fathers sub-samples. In Figure A3, we report the distributions of expected returns from different initial values in the whole sample and, as for beliefs about language development, for the mothers and fathers samples separately. As in Figure A2, we report KS tests for the difference between the two distributions and a test for the difference between the low and high initial conditions distributions.

Table A5: Measurement system for Beliefs

	A Measurement System for Beliefs					
	Couples sample		Mothers sample		Fathers sample	
	β	α	β	α	β	α
Language hard	1.000	0.000	1.000	0.000	1.000	0.000
High Dev.	-	-	-	-	-	-
Language hard	2.098	-0.023	1.220	0.024	1.314	-0.042
Low Dev.	*(0.514)*	*(0.201)*	*(0.257)*	*(0.124)*	*(0.243)*	*(0.114)*
Language medium	0.868	0.040	0.573	0.095	1.015	-0.090
High Dev.	*(0.175)*	*(0.074)*	*(0.136)*	*(0.066)*	*(0.138)*	*(0.065)*
Language medium	1.628	0.009	0.556	0.255	1.039	-0.066
Low Dev.	*(0.430)*	*(0.166)*	*(0.172)*	*(0.082)*	*(0.173)*	*(0.081)*
Language easy	0.746	0.065	0.057	0.296	0.877	-0.008
High Dev.	*(0.180)*	*(0.075)*	*(0.091)*	*(0.043)*	*(0.125)*	*(0.058)*
Language easy	1.082	0.167	-0.181	0.556	1.062	-0.075
Low Dev.	*(0.307)*	*(0.122)*	*(0.137)*	*(0.065)*	*(0.179)*	*(0.083)*
Socio-emotional Nine	0.332	0.156	0.566	0.216	0.965	-0.081
High Dev.	*(0.161)*	*(0.067)*	*(0.128)*	*(0.060)*	*(0.169)*	*(0.076)*
Socio-emotional Nine	0.397	0.259	1.420	0.072	1.168	-0.206
Low Dev.	*(0.317)*	*(0.133)*	*(0.238)*	*(0.106)*	*(0.199)*	*(0.090)*
Socio-emotional Five	0.362	0.128	0.379	0.190	0.819	-0.068
High Dev.	*(0.131)*	*(0.055)*	*(0.088)*	*(0.041)*	*(0.135)*	*(0.060)*
Socio-emotional Five	0.024	0.352	0.964	0.118	0.928	-0.182
Low Dev.	*(0.240)*	*(0.100)*	*(0.172)*	*(0.078)*	*(0.170)*	*(0.077)*
Socio-emotional Three	0.425	0.075	0.118	0.191	0.589	0.003
High Dev.	*(0.145)*	*(0.061)*	*(0.065)*	*(0.031)*	*(0.118)*	*(0.053)*
Socio-emotional Three	-0.180	0.374	0.368	0.168	0.702	-0.098
Low Dev.	*(0.202)*	*(0.084)*	*(0.117)*	*(0.055)*	*(0.170)*	*(0.078)*
Factor mean	0.346		0.303		0.348	
Factor variance	0.069		0.151		0.109	

Note: This table shows the loading factors (β) and the intercept (α) for each returns to parental investment on Language and Socio-emotional skills elicited in our survey which are estimated through a measurement system model. Columns (1) and (2) are for the couples sample, Columns (3) and (4) for the mothers sample, and Columns (5) and (6) for the fathers sample. Standard errors in parentheses. Source: Tz Pilot.

The results are similar to those of beliefs about the effect of parenting on cognitive development and language. In the whole sample, returns to parental investment in socio-emotional development are perceived to be higher for low than for high initial condi-

tions: the difference in means for the whole sample is equal to 0.044 (p-value=0.158). This effect comes from mothers, where the difference is equal to 0.098 (p-value=0.007). We see no such effect for the fathers sample, where the difference in means is equal to -0.063 (p-value=0.241).

Table A5 reports the loading factors (β) and the intercept (α) of the returns to parental investment on Language and Socio-emotional skills elicited in our survey.

A3.3 Measuring preferences

In Table A6, we report the share of the total additional resources allocated to each individual in the family (spouse, child, and self) for each of the three samples.

In Table A7, we report the shares allocated, for each sample, to the six different commodities. Table A8 shows the loading factors (β) and the intercept (α) for each ratio of the resources allocated to each spouse for four adult commodities (food, clothing, health, and transportation) from the allocation questions.

Table A6: Average share of expenditure allocated to household members

	Mother decision (s.e.)	Father decision (s.e.)	diff (p-value)	Couple decision (s.e.)	diff (p-value)
To self	0.268	0.257	-0.01	0.250	-0.02
	(0.008)	*(0.007)*	*(0.348)*	*(0.009)*	*(0.148)*
To spouse	0.175	0.219	0.04	0.232	0.06
	(0.009)	*(0.007)*	*(0.000)*	*(0.009)*	*(0.000)*
To child	0.558	0.524	-0.03	0.518	-0.04
	(0.011)	*(0.013)*	*(0.043)*	*(0.013)*	*(0.019)*

Note: This table shows the average share of expenditure to household members for the different samples. Standard errors are in parentheses. The p-values refer to the test of the difference between the mothers and fathers samples and the mothers and couples samples. Source: Tz Pilot.

Table A7: Allocation to the child

	Mother decision (s.e.)	Father decision (s.e.)	diff (p-value)	Couple decision (s.e.)	diff (p-value)
Clothing	6.628	5.559	-1.07	5.493	-1.13
	(0.225)	*(0.311)*	*(0.005)*	*(0.195)*	*(0.000)*
Food	6.062	5.338	-0.72	4.401	-1.66
	(0.302)	*(0.269)*	*(0.076)*	*(0.251)*	*(0.000)*
School exp.	7.434	7.529	0.09	7.282	-0.15
	(0.353)	*(0.573)*	*(0.886)*	*(0.456)*	*(0.791)*
Learning mat.	5.503	5.213	-0.29	5.697	0.19
	(0.247)	*(0.285)*	*(0.441)*	*(0.317)*	*(0.629)*
Health exp.	5.159	5.213	0.05	5.761	0.60
	(0.207)	*(0.252)*	*(0.866)*	*(0.232)*	*(0.054)*
Transportation	2.683	2.603	-0.08	2.430	-0.25
	(0.182)	*(0.202)*	*(0.769)*	*(0.199)*	*(0.349)*

Notes: This table shows the descriptive statistics of allocation of expenditure on children. Standard errors are in parentheses. The p-values refer to the test of the difference between the mothers and fathers samples and the mothers and couples samples. Source: Tz Pilot.

Table A8: Measurement system for relative taste for child human capital

	Relative taste for child human capital					
	Couples sample		Mothers sample		Fathers sample	
	β	α	β	α	β	α
Mother's clothing ratio	1.000	0.000	1.000	0.000	1.000	0.000
	-	-	-	-	-	-
Father's clothing ratio	1.009	-0.021	0.936	-0.039	0.729	0.007
	(0.058)	*(0.013)*	*(0.190)*	*(0.033)*	*(0.165)*	*(0.023)*
Mother's food ratio	0.850	0.043	4.294	-0.527	1.472	-0.031
	(0.172)	*(0.038)*	*(0.649)*	*(0.111)*	*(0.295)*	*(0.041)*
Father's food ratio	0.850	0.001	-0.011	0.111	1.229	-0.023
	(0.150)	*(0.032)*	*(0.179)*	*(0.031)*	*(0.252)*	*(0.036)*
Mother's health ratio	0.540	0.043	-0.028	0.115	1.410	-0.057
	(0.065)	*(0.014)*	*(0.106)*	*(0.018)*	*(0.294)*	*(0.041)*
Father's health ratio	0.530	0.030	-0.005	0.078	0.795	-0.006
	(0.060)	*(0.012)*	*(0.116)*	*(0.020)*	*(0.189)*	*(0.027)*
Mother's transportation ratio	0.540	0.043	-0.028	0.115	1.410	-0.057
	(0.065)	*(0.014)*	*(0.106)*	*(0.018)*	*(0.294)*	*(0.041)*
Father's transportation ratio	0.530	0.030	-0.005	0.078	0.795	-0.006
	(0.060)	*(0.012)*	*(0.116)*	*(0.020)*	*(0.189)*	*(0.027)*
Factor mean (variance)	0.175	(0.020)	0.162	(0.004)	0.135	(0.002)

Note: This table shows the loading factors (β) and the intercept (α) for each ratio of the resources allocated to each spouse for four adult commodities (food, clothing, health, and transportation) from the allocation questions which are estimated through a measurement system model. Standard errors in parentheses. Columns (1) and (2) are for the couples sample, Columns (3) and (4) for the mothers sample, and Columns (5) and (6) for the fathers sample. Source: Tz Pilot.

A3.4 Parental investment

In terms of the measure of parental investment, we use multiple sources of information:

Expenditures for children: Reported spending on children's health, clothing, shoes, toys, education, books and bedding.

Material investments: Reported frequency of acquiring the following products for their child: food, clothing, footwear, confectioneries medicine and oral rehydration salts, mosquito net, books, toys, and learning materials (notebooks, pens, and pencils).

Play material: The number of toyss the child has, made at home or bought, music instruments, books, and drawing equipment.

Adult activities with children: Reading books, singing, playing, and cooking with the child.

Didactic scale: From the Parental Style Questionnaire (PSQ): whether the primary caregiver (i) spends time playing with the child, (ii) provides the child with independent time to explore and learn on his/her own, (iii) provides the child with diverse social and interactive experience with same-age peers through play groups or informal get-together, (iv) provides the child with a structured organized, and predictable environment, (v) provides language learning opportunities for the child by labeling and describing qualities of objects, events or activities, reading books etc., (vi) provides the child with a variety of toys and objects for play and exploration, (vii) is patient with the child's misbehavior, and (viii) is flexible about behaviors expected from the child.

Social scale: The social scale from the PSQ accounts for whether the primary caregiver (i) promptly and appropriately responds to the child's expressed distress or discomfort, (ii) spends time talking to or conversing with child, (iii) provides child with quick and positive feedback to his/her bid for attention, (iv) provides child with affectionate displays of warmth and attention, and (iv) is aware of what child wants or feels.

Table A9: Parental Investment Components

	Couples sample	
	Mean	Observations
Parental Investment		
Raw Activity Score (/26)	14.89	142
Raw Material Investment Score (/12)	5.23	142
Raw Play Material Score (/8)	1.51	142
Social scale (/5)	4.28	142
Didactic scale (/8)	4.94	142
Expenditure on children share	0.32	142
Total Parental Investment (factor)	-0.38	142

Note: Mean and number of observations for markers of parental investment for the couples sample. In parenthesis the maximum value of each marker of parental investment. Source: Tz Pilot.

Table A10: A measurement system for parental investment

	Couples sample	
	β	α
Parental Activities	1.000	0.000
	-	-
Play Material	0.362	-0.201
	(0.136)	*(0.086)*
Material Investment	1.007	0.201
	(0.377)	*(0.176)*
Expenditure on	0.351	-0.080
children share	*(0.128)*	*(0.070)*
Social scale	0.657	0.181
	(0.329)	*(0.173)*
Didactic scale	0.415	0.031
	(0.235)	*(0.136)*
Factor mean (variance)	-0.376	(0.436)

Note: This table shows the loading factors (β) and the intercept (α) for markers of parental investment. Columns (1) and (2) are for the couples sample. Standard errors in parentheses. Source: Tz Pilot.

Table A9 reports descriptive statistics on the parental investment components and

Table A10 shows the loading factors (β) and the intercept (α) for markers of parental investment.

In Table A11, we report the results of regressing traditional measures of parental investment on log total expenditure as well as on beliefs and taste factors (and their interactions) and our measure of bargaining power.

Table A11: Modeling expenditure: Couples sample

	(1)	(2)	(3)	(4)	(5)
	Exp share	Exp share	Exp share	Exp share	Exp share
Log of total expenditure	-0.097***	-0.095***	-0.081***	-0.081***	-0.079***
	(0.030)	*(0.030)*	*(0.027)*	*(0.027)*	*(0.027)*
Relative taste for child human capital (RC)		-0.143	-0.136	-0.146	0.146
		(0.202)	*(0.179)*	*(0.330)*	*(0.421)*
Beliefs		-0.019	0.036	0.031	0.041
		(0.131)	*(0.121)*	*(0.198)*	*(0.122)*
Bargaining power (BP)			-0.009	-0.009	0.054
			(0.063)	*(0.064)*	*(0.105)*
Beliefs*RC				0.039	
				(1.117)	
BP*RC					-0.362
					(0.487)
R-squared	0.143	0.146	0.208	0.208	0.212
Observations	142	142	126	126	126

Note: The table displays regressions using the factor of expenditures share as a left hand side variable and the following right hand side variables: i) Log of total expenditure is the logarithm of total household expenditure from the household survey; ii) Bargaining power is a measure of female bargaining power, and more precisely, it is 1 minus the share that the woman is willing to pay to gain control over a fixed amount of money from our experiment; iii) Relative taste for child human capital is a factor of ratios of the resources allocated to each spouse for four adult commodities to the resources allocated to the child in our allocation experiment; and iv) Beliefs is a factor of the returns to parental investment on Language and Socio-emotional skills elicited in our survey. All models are controlled by child and mothers characteristics: child's age and gender, number of siblings in the household, dummy for mother's secondary education and mother's cognition skills measured by the Raven's test. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Source: Tz Pilot.

References

Acemoglu, D. and P. Restrepo (2019). "Automation and new tasks: How technology displaces and reinstates labor." *Journal of Economic Perspectives* **33**(2), 3–30.

Adams, A. and A. Andrew (2019). "Preferences and beliefs in the marriage market for young brides." Tech. rep., IFS Working Papers.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2019). "Missing growth from creative destruction." *American Economic Review* **109**(8), 2795–2822.

Alesina, A., M. Ferroni, and S. Stantcheva (2021). "Perceptions of racial gaps, their causes and ways to reduce them." Tech. Rep. 29245, NBER WP.

Alesina, A., A. Miano, and S. Stantcheva (2023). "Immigration and Redistribution." *The Review of Economic Studies* **90**(1), 1–39.

Alesina, A., S. Stantcheva, and E. Teso (2018). "Intergenerational mobility and preferences for redistribution." *American Economic Review* **108**(2), 521–54.

Almås, I. (2012). "International income inequality: Measuring ppp bias by estimating engel curves for food." *American Economic Review* **102**(2), 1093–1117.

Almås, I., A. Armand, O. Attanasio, and P. Carneiro (2018). "Measuring and changing control: Women's empowerment and targeted transfers." *The Economic Journal* **128**(612), F609–F639.

Almås, I., A. W. Cappelen, and B. Tungodden (2020). "Cutthroat capitalism versus cuddly socialism: Are americans more meritocratic and efficiency-seeking than scandinavians?" *Journal of Political Economy* **128**(5), 1753–1788.

Almås, I. and Å. A. Johnsen (2018). "The cost of a growth miracle–reassessing price and poverty trends in china." *Review of Economic Dynamics* **30**, 239–264.

Almås, I. (2012). "International income inequality: Measuring ppp bias by estimating engel curves for food." *American Economic Review* **102**(2), 1093–1117.

Ameriks, J., J. Briggs, A. Caplin, M. D. Shapiro, and C. Tonetti (2020). "Long-term-care utility and late-in-life saving." *Journal of Political Economy* **128**(6), 2375–2451.

Attanasio, O. and B. Augsburg (2016). "Subjective expectations and income processes in rural india." *Economica* **83**(331), 416–442.

Attanasio, O., B. Augsburg, and R. De Haas (2018). "Microcredit Contracts, Risk Diversification and Loan Take-Up." *Journal of the European Economic Association* **17**(6), 1797–1842.

Attanasio, O., A. Barr, J. C. Cardenas, G. Genicot, and C. Meghir (2012). "Risk pooling, risk preferences, and social networks." *American Economic Journal: Applied Economics* **4**(2), 134–67.

Attanasio, O. and V. Lechene (2014). "Efficient responses to targeted cash transfers." *Journal of Political Economy* **122**(1), 178 – 222.

Attanasio, O. and E. Pastorino (2020). "Nonlinear pricing in village economies." *Econometrica* **88**(1), 207–263.

Augsburg, B., O. P. Attanasio, M. Bornstein, J. Boyden, B. Caeyers, P. Espinoza, P. Jervis, S. Krutikova, E. Netsi, and A. Stein (2019). "Tanzania field pilot report." Tech. rep., The Institute for Fiscal Studies in collaboration with University of Oxford, NICHD and Economic Development Initiatives.

Ben-Akiva, M. E., D. McFadden, K. Train et al. (2019). *Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis*. Now.

Bernheim, B. D., D. Bjorkegren, J. Naecker, and M. Pollmann (2021). "Causal inference from hypothetical evaluations." Tech. rep., NBER, WP No 29616.

Bils, M. (2009). "Do Higher Prices for New Goods Reflect Quality Growth or Inflation?*." *The Quarterly Journal of Economics* **124**(2), 637–675.

Bils, M. and P. J. Klenow (2001). "Quantifying quality growth." *American Economic Review* **91**(4), 1006–1030.

Black, M. M., K. R. Bromley, V. Cavallera, J. Cuartas, T. Dua, I. Eekhout, G. Fink, M. J. Gladstone, K. L. Hepworth, M. Janus, P. K. Kariger, G. A. Lancaster, D. C. McCoy, G. P. J. McCray, A. Raikes, M. Rubio-Codina, S. van Buuren, M. Waldman, S. P. Walker, and A. M. Weber (2019). "The global scale for early development -gsed." *BMC Public Health* .

Blass, A. A., S. Lach, and C. F. Manski (2010). "Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability." *International Economic Review* **51**(2), 421–440.

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2019). "What drives differences in management practices?" *American Economic Review* **109**(5), 1648–83.

Bobba, M. and V. Frisancho (2020). "Self-perceptions about academic achievement: Evidence from mexico city." *Journal of Econometrics* .

Boneva, T. and C. Rauh (2018). "Parental beliefs about returns to educational investments—the later the better?" *Journal of the European Economic Association* **16**(6), 1669–1711.

Bornstein, M. H., W. A. Rothenberg, J. E. Lansford, R. H. Bradley, K. Deater-Deckard, A. Bizzego, and G. Esposito (2021). "Child Development in Low- and Middle-Income Countries." *Pediatrics* **148**(5). E2021053180.

Bourguignon, F., M. Browning, and P. A. Chiappori (2009). "Efficient intra-household allocations and distribution factors: Implications and identification." *The Review of Economic Studies* **76**(2), 503–528.

Browning, M. and P. A. Chiappori (1998). "Efficient intra-household allocations: A general characterization and empirical tests." *Econometrica* **66**(6), 1241–1278.

Buser, T., M. Niederle, and H. Oosterbeek (2014). "Gender, competitiveness, and career choices." *The quarterly journal of economics* **129**(3), 1409–1447.

Carson, R. T. (2012). "Contingent valuation: A practical alternative when prices aren't available." *Journal of Economic Perspectives* **26**(4), 27–42.

Cavatorta, E. and B. Groom (2020). "Does deterrence change preferences? evidence from a natural experiment." *European Economic Review* 103456.

Chamberlain, G. and Z. Griliches (1975). "Unobservables with a variance-components structure: Ability, schooling, and the economic success of brothers." *International Economic Review* 422–449.

Chen, X., L. P. Hansen, and P. G. Hansen (2020). "Robust identification of investor beliefs." Tech. rep., National Bureau of Economic Research.

Cherchye, L., P.-A. Chiappori, B. De Rock, C. Ringdal, and F. Vermeulen (2021). "Feed the children." Tech. rep., CEPR Discussion Paper No. DP16482.

Cherchye, L., B. D. Rock, and F. Vermeulen (2011). "The revealed preference approach to collective consumption behaviour: Testing and sharing rule recovery." *The Review of Economic Studies* **78**(1), 176–198.

Chiappori, P. A. (1988). "Rational household labor supply." *Econometrica* **56**(1), 63–90.

Chiappori, P.-A. (1992). "Collective labor supply and welfare." *Journal of Political Economy* **100**(3), 437–467.

Christensen, L. R., D. W. Jorgenson, and L. J. Lau (1975). "Transcendental logarithmic utility functions." *The American Economic Review* **65**(3), 367–383.

Costa, D. (2000). "American living standards, 1888-1994: evidence from consumer expenditures."

Crawford, I. and J. P. Neary (2021). "New Characteristics and Hedonic Price Index Numbers." *The Review of Economics and Statistics* 1–49.

Curtin, R. (2016). "Katona: a founder of behavioral economics." In "Handbook of Behavioral Economics," 30–47. Routledge.

Deaton, A. and J. Muellbauer (1980). "An almost ideal demand system." *The American economic review* **70**(3), 312–326.

Delavande, A. and B. Zafar (2019). "University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints." *Journal of Political Economy* **127**(5), 2343–2393.

Deming, D. J. (2017). "The Growing Importance of Social Skills in the Labor Market*." *The Quarterly Journal of Economics* **132**(4), 1593–1640.

Dizon-Ross, R. (2019). "Parents' beliefs about their children's academic ability: Implications for educational investments." *American Economic Review* **109**(8), 2728–65.

Dominitz, J. and C. F. Manski (1996). "Eliciting student expectations of the returns to schooling." *The Journal of Human Resources* **31**(1), 1–26.

Dominitz, J. and C. F. Manski (1997). "Using expectations data to study subjective income expectations." *Journal of the American Statistical Association* **92**(439), 855–867.

Dubois, P., R. Griffith, and M. O'Connell (2020). "How well targeted are soda taxes?" *American Economic Review* **110**(11), 3661–3704.

Dunbar, G. R., A. Lewbel, and K. Pendakur (2013). "Children's resources in collective households." *American Economic Review* **103**(1), 438–71.

Duncan, O. D. (1966). "Path analysis: Sociological examples." *American journal of Sociology* **72**(1), 1–16.

Einav, L., E. S. Leibtag, and A. Nevo (2008). "On the Accuracy of Nielsen Homescan Data." Economic Research Report 56490, US Dept. of Agriculture.

Engel, E. (1857). "Die productions-und consumtionsverhïltnisse des kïnigreichs sachsen." *Zeitschrift des Statistics Bï¿½reaus des Kï¿½niglich Sï¿½chsischen Ministeriums des Innern* **8**(1), 1–54.

Engel, E. (1895). *Die Lebenskosten belgischer Arbeiter-Familien früher und jetzt*. Heinrich.

Erdem, T., M. Keane, and T. Öncü (2005). "Learning about computers: An analysis of information search and technology choice." *Quantitative Market and Econonomics* **3**(3), 207–247.

Ertaç, S., A. Hortaçsu, and J. W. Roberts (2011). "Entry into auctions: An experimental analysis." *International Journal of Industrial Organization* **29**(2), 168–178.

Ferrario, B. and S. Stantcheva (2022). "Eliciting people's first-order concerns: Text analysis of open-ended survey questions." *AEA Papers and Proceedings* **112**, 163–69.

Gilbert, M., C. Clark, J. Stone, F. Perroux, D. Lieu, Evelpides, F. Divisia, Tinbergen, Kuznets, Smithies et al. (1949). "The measurement of national wealth: discussion." *Econometrica: Journal of the Econometric Society* 255–272.

Giustinelli, P., C. F. Manski, and F. Molinari (2019). "Precise or imprecise probabilities? evidence from survey response on late-onset dementia." Tech. rep., NBER.

Goldberger, A. S. (1971). "Econometrics and psychometrics: A survey of communalities." *Psychometrika* **36**(2), 83–107.

Goldberger, A. S. (1972). "Structural equation methods in the social sciences." *Econometrica* **40**(6), 979–1001.

Griffith, R. and M. O'Connell (2009). "The use of scanner data for research into nutrition*." *Fiscal Studies* **30**(3-4), 339–365.

Griliches, Z. (1974). "Errors in variables and other unobservables." *Econometrica:* **42**(6), 971–998.

GSED (2021). "The global scale for early development (gsed)." Tech. rep., Early Childhood Matters.

Guiso, L., P. Sapienza, and L. Zingales (2004). "The role of social capital in financial development." *American Economic Review* **94**(3), 526–556.

Guiso, L., P. Sapienza, and L. Zingales (2006). "Does culture affect economic outcomes?" *Journal of Economic Perspectives* **20**(2), 23–48.

Hamilton, B. W. (2001). "Using engel's law to estimate cpi bias." *American Economic Review* **91**(3), 619–630.

Harris, K. M. and M. P. Keane (1998). "A model of health plan choice:: Inferring preferences and perceptions from a combination of revealed preference and attitudinal data." *Journal of Econometrics* **89**(1-2), 131–157.

Hu, Y. and S. M. Schennach (2008). "Instrumental variable treatment of nonclassical measurement error models." *Econometrica* **76**(1), 195–216.

Jappelli, T. and L. Pistaferri (2000). "Using subjective income expectations to test for excess sensitivity of consumption to predicted income growth." *European Economic Review* **44**(2), 337–358.

Jayachandran, S., M. Biradavolou, and J. Cooper (2021). "Using machine learning and qualitative interviews to design a five-question survey module for women's agency." Tech. rep., Northwestern University Working Paper.

Jöreskog, K. G. and A. Goldberger (1975). "Estimation of a model with multiple indicators and multiple causes of a single latent variable." *Journal of the American Statistical Association* 631–639.

Jullien, B. (2000). "Participation constraints in adverse selection models." *Journal of Economic Theory* **93**(1), 1–47.

Juster, F. T. (1964). "Consumer sensitivity to the price of credit." *The Journal of Finance* **19**(2), 222–233.

Juster, F. T. (1966). "Consumer buying intentions and purchase probability: An experiment in survey design." *J. of the American Statistical Association* **61**(315), 658–696.

Juster, F. T. and R. P. Shay (1964). *Consumer Sensitivity to Finance Rates: An Empirical and Analytical Investigation*. NBER.

Kaufmann, K. and L. Pistaferri (2009). "Disentangling insurance and information in intertemporal consumption choices." *American Economic Review* **99**(2), 387–92.

Kesternich, I., F. Heiss, D. McFadden, and J. Winter (2013). "Suit the action to the word, the word to the action: Hypothetical choices and real decisions in medicare part d." *Journal of Health Economics* **32**(6), 1313–1324.

Keynes, J. M. (1936). "The general theory of interest, employment and money."

Kuznets, S. (1941). "Statistics and economic history." *The Journal of Economic History* **1**(1), 26–41.

Kuznets, S. et al. (1937). "National income and capital formation, 1919-1935." *NBER Books* .

Lechene, V., K. Pendakur, and A. Wolf (2022). "Ols estimation of the intra-household distribution of expenditure." *Journal of Political Economy* **130**, *forthcoming*.

Lee, J., R. Wedow, A. Okbay, E. Kong, O. Maghzian, M. Zacher, T. Nguyen-Viet, P. Bowers, J. Sidorenko, R. Linnér, M. Fontana, T. Kundu, C. Lee, H. Li, R. Li, R. Royer, P. Timshel, R. Walters, E. Willoughby, and D. Cesarini (2018). "Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals." *Nature Genetics* **50**.

List, J. A. and C. A. Gallet (2001). "What experimental protocol influence disparities between actual and hypothetical stated values?" *Environmental and Resource Economics* **20**(3), 241–254.

List, J. A., J. Pernaudet, and D. Suskind (2021). "It all starts with beliefs: Addressing the roots of educational inequities by shifting parental beliefs." WP 29394, NBER.

Louviere, J., D. Hensher, and J. Swait (2000). *Stated Choice Models: Analysis and Application*. Cambridge University Press.

Luce, R. and P. Suppes (1965). "Preference, utility, and subjective utility." *Handbook of Mathematical Psychology, III, New York: Wiley* 249–409.

Luce, R. D. (1959). *Choice Behavior. A Theoretical Analysis*. New York: Wiley.

Luce, R. D. and J. W. Tukey (1964). "Simultaneous conjoint measurement: A new type of fundamental measurement." *Journal of mathematical psychology* **1**(1), 1–27.

Maskin, E. and J. Riley (1984). "Monopoly with incomplete information." *RAND Journal of Economics* **15**(2), 171–196.

McCoy, D. C., J. Seiden, M. Waldman, and G. Fink (2021). "Measuring early childhood development." *Annals of the New York Academy of Sciences* **1492**(1), 3–10.

Miller, G., Á. de Paula, and C. Valente (2020). "Subjective expectations and demand for contraception." Working Paper 27271, National Bureau of Economic Research.

Murphy, P. G., James J.and Allen, T. H. Stevens, and D. Weatherhead (2005). "A meta-analysis of hypothetical bias in stated preference valuation." *Environmental and Resource Economics* **30**(3), 313–325.

Neary, J. P. (2004). "Rationalizing the penn world table: True multilateral indices for international comparisons of real income." *American Economic Review* **94**(5), 1411–1428.

Nordhaus, W. D. (1998). "Quality change in price indexes." *Journal of Economic Perspectives* **12**(1), 59–68.

Paiella, M. and L. Pistaferri (2017). "Decomposing the Wealth Effect on Consumption." *The Review of Economics and Statistics* **99**(4), 710–721.

Pistaferri, L. (2001). "Superior Information, Income Shocks, and the Permanent Income Hypothesis." *The Review of Economics and Statistics* **83**(3), 465–476.

Pistaferri, L. (2003). "Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply." *Journal of Labor Economics* **21**(3), 729–754.

Potter, S., M. Del Negro, G. Topa, and W. Van der Klaauw (2017). "The advantages of probabilistic survey questions." *Review of Economic Analysis* **9**(1), 1–32.

Salz, T. and E. Vespa (2020). "Estimating dynamic games of oligopolistic competition." *The RAND Journal of Economics* **51**(2), 447–469.

Schennach, S. M. (2004). "Estimation of nonlinear models with measurement error." *Econometrica* **72**(1), 33–75.

Scur, D., R. Sadun, J. Van Reenen, R. Lemos, and N. Bloom (2021). "The world management survey at 18." *Oxford Review of Economic Policy* **37**(2), 231–258.

Stantcheva, S. (2022). "Understanding trade." Tech. Rep. 30040, NBER.

Stone, R. (1954). "Linear expenditure systems and demand analysis: an application to the pattern of british demand." *The Economic Journal* **64**(255), 511–527.

Stone, R. (1984). "Richard stone-prize lecture: The accounts of society." *Nobelprize. org. Nobel Media AB* .

Van der Klaauw, W. and K. I. Wolpin (2008). "Social security and the retirement and savings behavior of low-income households." *J. of Econometrics* **145**(1-2), 21–42.

Visco, I. (1984). *Price Expectations in Rising Inflation*. Contributions to economic analysis. North-Holland.

Wiswall, M. and B. Zafar (2015). "Determinants of college major choice: Identification using an information experiment." *The Review of Economic Studies* **82**(2), 791–824.

Wright, S. (1934). "The method of path coefficients." *The annals of mathematical statistics* **5**(3), 161–215.