# New measures for richer theories:
## some thoughts and an example*

Orazio Attanasio† Víctor Sancibrián‡ Federica Ambrosio §

December 2025

**Abstract**

For a long time, the majority of economists doing empirical work relied on choice data, while data based on answers to hypothetical questions, stated preferences or measures of subjective beliefs were met with some skepticism. Although this has changed recently, much work needs to be done. In this paper, we emphasize the identifying content of new economic measures. In the first part of the paper, we discuss where the literature on measures in economics stands at the moment. We first consider how the design and use of new measures can help identify causal links and structural parameters under weaker assumptions than those required by approaches based exclusively on choice data. We then discuss how the availability of new measures can allow the study of richer models of human behavior that incorporate a wide set of factors. In the second part of the paper, we illustrate these issues with an application to the study of risk sharing and of deviations from perfect risk sharing.

*This paper develops some of the ideas in Almås, Attanasio, and Jervis (2024). We have greatly benefited from comments and discussion by Arnaud Maurel. We also thank Manuel Arellano, Margherita Borella, Flavio Cunha, Mariacristina De Nardi and Aureo De Paula. Attanasio is grateful for the hospitality provided by the Collegio Carlo Alberto in Torino, EIEF in Roma and Università Federico II in Napoli. Sancibrián acknowledges financial support from Grant PRE2022-000906 funded by MCIN/AEI/10.13039/501100011033 and by "ESF +", the Maria de Maeztu Unit of Excellence CEMFI MDM-2016-0684, funded by MCIN/AEI/10.13039/501100011033, Fundación Ramón Areces and CEMFI. A big part of this work was done while Sancibrián was at CEMFI. Edited by Magne Mogstad.

†Yale University, MeToD @UNICAL, CEPR and NBER: orazio.attanasio@yale.edu

‡Bocconi University and IGIER: victor.sancibrian@unibocconi.it

§Università di Napoli Federico II: federica.ambrosio@unina.it

# 1 Introduction

In this paper, we discuss how new measurement tools can be used in economics to allow more robust empirical analyses and, indirectly, much richer theoretical models. Building on the arguments in Almås, Attanasio, and Jervis (2024), we discuss new approaches to measurement that, informed by the theoretical models of interest, can yield identification of structural parameters under a weaker set of assumptions.

Measurement tools inspired by the need to identify certain structural parameters can facilitate the analysis of complex models with relatively simple econometric techniques. Furthermore, when econometric tests reject benchmark structural models, extensions of such theories can be analyzed and made empirically salient with the help of specific measures that capture factors likely to inform human behavior. This approach may bring about a shift in empirical work in economics. While, until recently, most of the work used exclusively choice data and a "revealed preferences" approach, in recent years researchers have started to use data on subjective expectations, personal beliefs, attitudes, and social norms. The construction of such data is not trivial, which might have been one of the reasons why economists have been skeptical in their use for a long time. These difficulties make the validation of new measures necessary. In this sense, the combined use of choice data and alternative measures is very important.

The emphasis in the first part of the paper (Section 2) is on identification issues, considered from two different angles. First, we stress how using alternative measurement tools changes and simplifies the identification problem. By designing new measures, dynamic expectations models or models that involve difficult selection problems can be identified with much weaker assumptions. Second, we argue that the availability of measures capturing different latent factors and constructs can allow the empirical estimation of more sophisticated and nuanced structural models. What we present in the first part is not a comprehensive survey, such as Koşar and O'Dea (2023), both because we do not focus only on subjective expectations and, more importantly, because our main message is how new measures, when based on the needs of economic theory, can make possible the empirical characterization of a wide class of rich theoretical models of human behavior.

In the second part of the paper (Sections 3 to 5), we look at a model of dynamic resource allocations to illustrate the themes discussed in the first part. In particular, we consider models of insurance of idiosyncratic income shocks and analyze them using novel measures collected in a survey from Colombia.

Numerous studies have examined the insurance of idiosyncratic income shocks at various levels. In his seminal work, Townsend (1994) analyzed risk sharing among individuals in three Indian villages and introduced a novel method to test for perfect insurance. This approach has since been widely applied in different contexts. Most studies reject perfect risk sharing. However, rejections of such benchmark complete-market models may be of limited use if no additional information is provided on the mechanisms behind their empirical failure: *why are markets incomplete?*

One possible approach is to consider alternative measurement tools to improve our understanding of risk-sharing arrangements. We do so by modeling deviations from first-best risk sharing. We start by measuring the deviations from the first-best à la Townsend (1994) by estimating the degree to which idiosyncratic resource shocks are reflected in idiosyncratic consumption changes within a predefined risk-sharing group. We then consider three alternative theoretical models that could prevent perfect insurance and use novel measurements to assess their plausibility. We first consider the possibility that, to be sustained, perfect risk-sharing needs some form of social cohesion. Second, we analyze the possibility that perfect risk sharing within a certain group is prevented by information problems. Third, we look at models of imperfect enforceability of contracts that relate the amount of risk sharing to the properties of perceived income processes, such as the variance and persistence of household income.

We implement these ideas using a rich dataset that was collected several years ago in Colombia to evaluate the impacts of a conditional cash transfer program. Our results show that deviations from first-best risk sharing seem to be related to the quality of information available in the risk-sharing group considered and to the properties of the income processes available to the individuals in the risk-sharing groups. In particular, consistent with imperfect information models, better risk sharing is observed in groups with better quality of information. At the same time, and consistent with some models of imperfect enforceability of contracts, risk sharing is also better in villages with relatively high risk and low persistence of idiosyncratic income. Our models of social cohesion, on the other hand, do not seem to be related to the extent of risk sharing.

In summary, in the second part of the paper we start by providing a framework to quantify deviations from perfect risk sharing and discuss models that could relate these deviations to frictions related to the three sets of variables we will be using: social capital and cohesion, information quality, and properties of the income process. In Section 4, we describe the data and the novel measures we use. Finally, in Section 5, we present the empirical results of the analysis. Additional discussion and results are relegated to the Supplemental Appendix.[1]

[1]The Supplemental Appendix can be accessed at `https://sancibrian-v.github.io/files/new_measures-supp.pdf`.

## 2 New measures and identification

For several decades, economists — with some exceptions — have been skeptical about using data on subjective beliefs and expectations, stated preferences, and intentions. Instead, the emphasis has been on a *revealed preferences* approach: meaningful data are those that refer to choices agents make in response to real incentives, resources, and prices.

The exclusive use of choice data imposes important limitations on empirical work. In particular, endogeneity and selection problems become paramount: agents make decisions based on a range of variables, many of which are unobservable and, in all likelihood, related to those included in the choice model under study. Identification, then, requires strong and often untestable assumptions.

In recent years, the profession has increasingly adopted innovative, researcher-designed measures—most notably in the elicitation of subjective expectations, as pioneered by Manski (2004) and later reviewed by Hurd (2009). Recent work has deepened this approach: for example, Giustinelli, Manski, and Molinari (2019) and Giustinelli, Manski, and Molinari (2022) explore how respondents may be uncertain about their own probability assessments. Some recent contributions go beyond traditional elicitation methods. Stantcheva (2022) advocates for survey designs that reveal behavioral mechanisms through structured variation, while Almås, Attanasio, and Jervis (2024) argue for aligning measurement tools with theoretical models to better identify and estimate more realistic behavioral frameworks. We build on this perspective by elaborating on two key directions.

First, we argue that well-designed, validated measurement tools can help identify structural parameters under weaker restrictions than those required by approaches based exclusively on choice data and revealed preference methods. While these tools are not immune to endogeneity concerns, stated preferences and hypothetical scenarios can help mitigate such issues. Second, when standard models are rejected empirically, innovative measures enable estimation of richer models that better reflect real-world frictions. Our empirical application in the second part of the paper illustrates this approach. Finally, given the complexity of designing these tools, rigorous validation is essential to ensure their reliability and interpretability.

### 2.1 Theory and measures: a conceptual framework

In many econometric applications, the main goal is to identify the parameters of a function linking two or more theoretical constructs from economic models. Such a function can be written as

$$f(\theta_{it}, v_{it}, \varepsilon_{it}) = 0, \tag{1}$$

where $\theta_{it} = \{\theta_{it}^{(1)}, \theta_{it}^{(2)}, \theta_{it}^{(3)}\}$ might be observable variables (possibly vectors), and $\varepsilon_{it}$ and $v_{it}$ are unobservable; the subscripts $i$ and $t$ refer to individuals and time. The difference between $\varepsilon_{it}$ and $v_{it}$ is that $v_{it}$ might be related to the observable variables, whereas $\varepsilon_{it}$ is not. The objects of interest are the relationships between $\theta_{it}^{(1)}$ and $\theta_{it}^{(2)}$, while those between $\theta_{it}^{(1)}$ and $\theta_{it}^{(3)}$ are not of particular interest.

The function $f$ can be nonlinear and capture the fact that some variables are only observed for a selected sample or driven by binary choices related to the questions of interest. For example, wages or earnings are observed only for individuals who are employed and may be influenced by job transitions. In addition to (1), for expositional simplicity, we consider a linear version of $f$:

$$\theta_{it}^{(1)} = \alpha_2 \theta_{it}^{(2)} + \alpha_3 \theta_{it}^{(3)} + v_{it} + \varepsilon_{it}, \tag{2}$$

where the presence of selection issues could nonetheless make the properties of the residuals very complex and induce strong dependencies between observables and unobservables.

Economic models are often populated by abstract theoretical constructs. In many situations, the theoretical constructs of interest in equations (1) or (2) (that is, the variables $\theta_{it}$) are not directly observable. However, they might be connected to a set of *markers* or observable variables. The connection between the available measures and the latent variables of interest can be formalized via a *measurement system*, as in Cunha, Heckman, and Schennach (2010). These ideas have been discussed in economics and other disciplines for many years; see, for instance, Goldberger (1971, 1972); Goldberger and Duncan (1973); Goldberger and Jöreskog (1975).

A widely used specification of such measurement systems is the following:

$$m_{it}^{(k_j)} = \Lambda_{0t}^{(k_j)} + \Lambda_{1t}^{(k_j)} \theta_{it}^{(j)} + u_{it}^{(k_j)}, \quad k_j = 1, \ldots, K^{(j)}, \ j = 1, 2, 3. \tag{3}$$

where $m_{it}^{(k_j)}$ in equation (3) are the *markers* of the factors of interest $\theta_{it}^{(j)}$. When the matrices $\Lambda_{1t}^{(k_j)}$ are diagonal, the system is described as *dedicated*, meaning each measure loads on only one latent factor. The variables $u_{it}^{(k_j)}$ represent *measurement error* that prevents the latent variables of interest $\theta_{it}$ from being observable. If a sufficient number of markers is available, the parameters of equation (3) can be identified under assumptions about the measurement error—chiefly, *independence* across measures (Cunha et al., 2010; Kotlarski, 1967). While intuitive, this assumption (and its need) should be kept in mind when designing questionnaires and implementing data collections, as it can inform protocols that can guarantee that it is satisfied.

When working with a measurement system, several additional issues must be addressed. A central concern in many applications is establishing the appropriate metric for the latent variables, which also determines how different measures are aggregated. This often involves normalizing the $\Lambda$ parameters and is especially important when comparing latent variables across contexts or time periods to ensure consistency and comparability.

The main econometric challenge in identifying parameters in models like equation (2) stems from the unobserved vectors $\varepsilon_{it}$ and $v_{it}$. The use of markers or latent factor estimates in estimating correlations or regressions typically leads to biased results. Two key points deserve emphasis: (i) the structure of the model and, crucially, the nature of the available data shape the properties of the residuals $\varepsilon_{it}$ and $v_{it}$, as well as viable identification strategies; (ii) addressing these challenges may require innovative data collection efforts and the development of tailored measurement tools.

Some examples help clarify these issues and illustrate how innovative measures can be crafted to address identification challenges and answer specific research questions.

**Data collection strategies to deal with measurement error.** Most microeconomic data, with the possible exception of administrative data, are affected by measurement error. An explicit consideration of the presence of such noise can be very important in terms of the econometric approach used. In addition, it can provide key insights into the design of data collection strategies.

Suppose, for example, that a theoretical construct $\theta_{it}^{*(2)}$ is observed with (additive) measurement error, so that the observable variable $\theta_{it}^{(2)}$ in equation (2) is related to one of the components of the unobservable variables. It is well-known that, under certain regularity conditions, the full distribution of $\theta_{it}^{*(2)}$ is identified from two repeated measurements with independent noise (Kotlarski, 1967).[2] As an application of these ideas in the context of survey design, Browning and Crossley (2009) point out that it might be better to have two noisy measures of consumption rather than trying to improve the reliability of a single one. A repeated-measurements framework is a useful approach, since it naturally allows for noise in the elicitation of these novel measures (see for instance the review by Schennach, 2016) . We return to this issue in Section 2.3.

A similar logic applies to strategies for addressing non-response bias. Respondents who opt out of long surveys are often a non-random sample and may differ systematically in key variables of interest. Correcting for this requires identifying factors that affect survey participation without directly influencing the outcome of interest. One effective approach is to randomize aspects of data collection—such as interviewer assignment or financial incentives—to generate exogenous variation in response rates. While this contrasts with typical practices that match

[2]Given $\theta_{it}^{(2)} = (\theta_{1it}^{(2)}, \theta_{2it}^{(2)})'$ with $\theta_{jit}^{(2)} = \theta_{it}^{*(2)} + e_{jit}$ for $j = 1, 2$, Kotlarski's lemma says that if $(\theta_{it}^{*(2)}, e_{1it}, e_{2it})$ are mutually independent with non-vanishing characteristic functions, the distribution of $\theta_{it}^{*(2)}$ (and that of the errors) is identified given that of $\theta_{it}^{(2)}$ up to location.

skilled interviewers to challenging cases, it can yield valuable information for correcting selection bias.

A recent example in this direction is the work by Dutz, Huitfeldt, Lacouture, Mogstad, Torgovitsky, and van Dijk (2025), who exploit the randomized allocation of financial incentives in a survey on the labor market consequences of the COVID-19 lockdown in Norway to test and account for nonresponse bias.

**Evaluating social programs: RCTs vs non-experimental variation.** Economists have long been interested in evaluating the impact of specific interventions — such as the negative tax experiments performed in the 1960s or training programs — on economic outcomes — such as individual earnings or income. A major challenge in this context is that participation in a program, for instance a job training initiative, may be influenced by factors that independently affect the outcome of interest, like innate ability. When relying exclusively on choice data to compare participants and non-participants, researchers must identify variables that influence program participation but not the outcome of interest. Consequently, in many cases, identification of the parameters of interest relies on assumptions that may be arbitrary.

The use of randomized controlled trials (RCTs) constitutes a useful approach to the identification of *some* of the parameters of interest, as participation in the program is randomly assigned and, therefore, is independent of the outcome of interest. In the spirit of this paper, one can interpret the promotion of RCTs to evaluate social programs, as, for example, in Duflo, Glennerster, and Kremer (2007) and Athey and Imbens (2017) as promotion of new *measurement tools* that build *counterfactual scenarios* as real choices. In other words, an RCT could be interpreted as a measurement tool, as it makes available observations from an environment that has been manipulated by the researcher. It provides information on the average behavior of individuals in a counterfactual context, which can be used to identify and estimate a variety of parameters of interest, starting with the impact of an intervention. When RCTs are available or can be constructed, approaches that rely on untestable assumptions, such as instrumental variables or difference-in-differences, could be avoided.

It is important to emphasize, as discussed in Wolpin (2013), that even when an RCT introduces exogenous variation in a variable of interest, this does not guarantee that all structural parameters can be identified. While one can estimate the reduced-form effect of changes in $\theta_{it}^{(2)}$ on $\theta_{it}^{(1)}$, establishing a causal interpretation often requires unpacking the mechanisms through which $\theta_{it}^{(2)}$ influences other variables that also affect $\theta_{it}^{(1)}$. Nevertheless, variation induced by an RCT can be leveraged to test and identify richer structural models.[3]

[3] See also Todd and Wolpin (2006) and Attanasio, Meghir, and Santiago (2012b).

Consider equation (2), where the goal is to identify $\alpha_2$, the causal effect of $\theta_{it}^{(2)}$ on $\theta_{it}^{(1)}$, holding other variables constant. If an RCT shifts $\theta_{it}^{(2)}$ but also indirectly affects $\theta_{it}^{(3)}$, then a simple regression of $\theta_{it}^{(1)}$ on $\theta_{it}^{(2)}$ captures only a reduced-form effect. Identifying $\alpha_2$ requires understanding the induced relationship between $\theta_{it}^{(2)}$ and $\theta_{it}^{(3)}$—something RCTs can also help uncover. Ultimately, the ability to recover structural parameters depends on the theoretical model guiding the analysis.

**Dynamics in panel data models.** Longitudinal data provide economists with the possibility of characterizing the dynamic properties of individual income processes. Quantifying income volatility and persistence is of key importance when modeling individual choices of consumption, saving, labor supply, and many other variables. A number of important contributions started looking at the dynamic properties of earnings and income data, including Lillard and Willis (1978), Lillard and Weiss (1979), MaCurdy (1982), and Abowd and Card (1989).

A key issue is the extent to which the persistence observed in individual earnings and income is driven by permanent differences among individuals, such as innate ability, or by genuine state-dependence. The canonical earnings process considered in this context can be expressed as

$$y_{i,t+1} = \rho y_{i,t} + \mu_i + \sigma\varepsilon_{i,t+1}, \quad \text{for } i = 1,\ldots,N; t = 1,\ldots,T, \tag{4}$$

where $y_{i,t+1}$ is log earnings of individual $i$ at time $t$, $\mu_i$ captures permanent unobserved heterogeneity and $\varepsilon_{i,t+1}$ are transitory innovations. Equation (4) can be seen as a special case of equation (2), with $x_{i,t} = y_{i,t-1}$ and $\mu_i$ being part of the unobserved components $v_{it}$. The presence of unit fixed effects can cause systematic biases, as observed by Nickell (1981). Since $y_{i,t}$ is not strictly exogenous, OLS and within-groups estimators are generally biased in short panels (large-$N$, small-$T$).

An influential contribution in this area is Arellano and Bond (1991), which introduced a generalized method of moments (GMM) estimator to address bias in dynamic panel models. Their first-differenced GMM approach eliminates fixed effects and uses lagged levels of the dependent variable as instruments in equation (4). While widely adopted, this method relies on strong assumptions about the income process, including log-linearity and additive fixed effects.[4]

An alternative approach to the identification of flexible income processes, based on subjective income expectations, is taken in Arellano, Attanasio, Crossman, and Sancibrián (2024), which has two advantages. First, it is directly informative about the income process as *perceived* by individuals, who presumably use these perceptions to make relevant decisions. Second, it can identify the parameters of interest, such as persistence, from a very short longitudinal dimension.[5]

[4]See Arellano, Blundell, and Bonhomme (2017) for recent work on nonlinear models of earnings dynamics.

[5]In the absence of fixed effects, the dynamic properties of the model can be identified from a single cross section.

The key to this lies in the use of *subjective expectations* of future income as outcomes of interest *in conjunction with current income realizations*. This distinction changes the nature of unobserved disturbances in these models, which contain only measurement errors in the elicited probabilities, and not future shocks.[6]

The validity and credibility of this approach, when using subjective expectations, is based on the quality of the data. As we discuss later, while much progress has been made in the design and use of subjective expectations data since they were strongly advocated by Manski (2004), several open issues remain in how they are elicited and collected. As Manski (1990) puts it in a different context, the issue is not *what* is collected, but *how*.

**Selection models and their estimation with counterfactual data.** Selection is a central concern in many economic models. When researchers rely exclusively on observed choices—such as enrolling in education, entering the labor market, or switching jobs—they only observe outcomes for individuals who self-select into particular alternatives. Because these decisions are often influenced by variables correlated with the outcomes of interest, making causal inference difficult.

New measures capturing the determinants of individual decisions, like the expected returns of specific investments, can help address this challenge. However, these variables may themselves be endogenous, subject to the same unobserved factors that affect the choices themselves. For example, Wiswall and Zafar (2015) examine how expectations of future earnings influence college major choice. Recognizing the potential endogeneity of beliefs, they experimentally vary the information provided to participants to identify causal effects.

A different approach is made possible by the availability of data on choices in hypothetical, counterfactual scenarios, which can greatly facilitate the identification and estimation of structural models. Arcidiacono, Hotz, Maurel, and Romano (2020), Giustinelli and Shapiro (2024) and Wiswall and Zafar (2021) elicit counterfactual potential outcomes in order to recover ex-ante treatment effects, which allows them to directly identify the causal links of interest.

Well-designed measures can also be instrumental in identifying components of structural models that are either not identifiable or identifiable only under strong assumptions when using observed choice data alone. Pantano and Zheng (2013), Meango (2023), and Wiswall and Zafar (2018) demonstrate how information on subjective probabilities of discrete hypothetical choices can be used to identify moments of permanent unobserved heterogeneity.[7]

With fixed effects, $T = 2$ is sufficient and the within-groups estimator does not suffer from Nickell bias.

[6] Moreover, the approach in Arellano et al. (2024) can accommodate flexible nonlinear extensions such as those discussed in Arellano et al. (2017), and even augmented versions allowing for fixed effects.

[7] Due to the discrete nature of the data, point identification typically requires distributional or functional form assumptions. We thank Arnaud Maurel for emphasizing this point.

Another example is the complex structural model in Altonji, Smith, and Vidangos (2013), where individual earnings depend on two latent factors reflecting productivity and job match quality, as well as an individual "ability" and a "mobility" fixed effect. Individuals move in and out of unemployment and when employed, they decide whether to stay in their current job or change jobs. The authors identify and estimate their complex model using Simulated Method of Moments and PSID data, using some strong assumptions, particularly in dealing with selection and heterogeneity.

This example illustrates the challenges discussed in connection with equation (2): even when assuming a simple (log-)linear relationship between current and lagged latent factors, selection leads to highly nonlinear relationships among observables. In a recent piece, Arellano, Attanasio, Borella, De Nardi, and Paz-Pardo (2025) estimate such a model using subjective expectations data on future earnings from the New York Fed's Survey of Consumer Expectations. This rich dataset includes beliefs about the distribution of future job offers, the likelihood of receiving them, and the probability of accepting them. Crucially, Arellano et al. (2025) exploit questions about potential offers — distinct from expected future earnings — to identify and estimate model parameters using linear within estimators and GMM. Similarly, Arcidiacono et al. (2020) combine expectations data with counterfactual scenarios to estimate a structural model of occupational choice.

Another study that uses subjective expectations and hypothetical choice data is Briggs, Caplin, Leth-Petersen, and Tonetti (2024), where the authors discuss identification of marginal treatment effects of the type we mentioned above when discussing the use of RCTs. They leverage subjective expectations data to estimate a model of program participation and address endogeneity issues.[8]

Importantly for our discussion, these studies show how data on how individuals view a variety of future events and scenarios can address the endogeneity and selection issues that are relevant when using choice data. Information on the (perceived) probability distributions that individuals use to make decisions can be much richer than the limited binary information on actual choices.

We emphasize that in many contexts, the use of new measures can be extremely advantageous when made *in connection with* choice data. And even when identification requires the latter, the former could provide more precision in the estimation of the structural parameters, a point made, for example, by van der Klaauw (2012), Van der Klaauw and Wolpin (2008), and Wolpin (1999).

[8] Méango and Poinas (2023) use subjective expectations data to identify a generalized Roy model they use to study the decision to overstay of asylum seekers in Germany.

### 2.2 Measures to identify richer models

So far, we have argued that innovative measures — ranging from those implicitly derived from RCTs to information on subjective expectations and hypothetical counterfactuals — can help solve some of the identification problems relevant in many econometric applications. In the literature, some important contributions have pointed to the use of innovative measures. Lancaster and Chesher (1983, p. 1661), for instance, advocate the use of "the answers to two simple questions" which could be interpreted as providing information about the distribution of offer wages and reservation wages to "*deduce* structural parameters rather than estimate them".

More generally, the availability of innovative measures can be used to study empirically much richer models than those studied with choice data, and similarly the design of those measures is driven by the needs of rich theoretical models. In this section, we discuss several novel measures that have been used in the recent literature. Our empirical application is yet another illustration of the opportunities the availability of novel measures can afford.

**Measuring preferences and norms.** A growing body of literature recognizes that economic behavior and choices are driven by various factors, some of which are not easily observed. In many contexts, direct measures of preferences can be very useful. Analogously, other variables, such as societal norms and conventions, can play a large role in explaining individual behavior. Many studies have shown substantial heterogeneity in tastes and preferences among individuals.

An early study discussing measures of different dimensions of preferences is Barsky, Juster, Kimball, and Shapiro (1997). The authors discuss an innovative set of measures, stating that "participants in the Health and Retirement Study were asked to respond to hypothetical situations specifically designed to elicit information about their risk aversion, subjective rate of time preference, and willingness to substitute" (p. 538). In addition to a detailed and careful discussion of the way the measures of risk aversion are constructed, they also relate them to "risky behaviors", which serves as an important validation exercise.[9]

Risk attitudes are not the only aspect of preferences that have been measured, nor the only ones whose measures have been validated. Frederick, Loewenstein, and O'Donoghue (2002) relate estimates of patience to saving behavior, while Fisman, Kariv, and Markovits (2007) and DellaVigna, List, and Malmendier (2012) relate measures of altruism to charitable giving.

More generally, social norms and social capital are clearly important elements of how human

[9]Other studies measuring risk attitudes in large-scale surveys include Guiso and Paiella (2008) and Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner (2011). Both articles validate the measures of risk attitudes by relating them to observed choices and demonstrating a substantial amount of heterogeneity in observed preferences.

societies work. An increasing number of papers attempt to use direct measures of these factors. Examples using data from different parts of the world include Alesina and La Ferrara (2002), Guiso, Sapienza, and Zingales (2004) and Field, Pande, Rigol, Schaner, and Troyer Moore (2021). Related to measures of social norms are studies that look at different attitudes individuals have towards a variety of topics, including perceptions of fairness, such as Almås, Cappelen, and Tungodden (2020b); Almås, Cappelen, Sørensen, and Tungodden (2020a). When looking at participation in the stock market, Guiso, Sapienza, and Zingales (2008) show the importance of trust in institutions and market transparency using novel Dutch data on these variables.

**Choices within the household.** When looking at the allocation of resources within the family, a model that has received considerable attention is the so-called collective model, proposed by Chiappori (1988, 1992). The collective framework or other models of resource allocation mechanisms within the household constitute a good example of a situation where new measures can yield important new insights. Much important empirical work has been done on these models using household expenditure data, such as Bourguignon, Browning, and Chiappori (2009). However, a challenge in the analysis of these models is the scarce availability of data on consumption at the individual level; most data sets contain information on expenditure at the household level.

Some recent papers, mentioned in Supplemental Appendix A, have derived clever tests of models of the family using only household expenditure data, often exploiting the fact that some commodities can be defined as assignable, while others can be identified as public goods within the household. However, to identify some of the key features of resource allocation, these studies impose strong restrictions on preferences, ranging from the presence of assignable goods not affecting the utility of some household members to limitations on the heterogeneity in tastes across different households.

While additional data on individual consumption would undoubtedly provide richer insights into the intra-household allocation of resources, it is not the only type of information that can enable the estimation of more nuanced models of individual behavior. Ashraf (2009) demonstrates that spouses' decisions in experimental settings vary depending on the extent of information and control each partner has over household finances. Almås, Armand, Attanasio, and Carneiro (2018) develop a measure of intra-household bargaining power in the context of evaluating targeted conditional cash transfers in Macedonia. Their incentivized experiment reveals that the willingness to pay for control over resources is influenced by who in the household is targeted by the transfer.[10]

[10] Almås et al. (2024) show that this measure of bargaining power is predictive of actual household decisions in Tanzania. Jayachandran, Biradavolou, and Cooper (2021) find that in other contexts, alternative indicators of relative power within couples may better predict observed choices than the approach proposed by Almås et al. (2018).

Households make many different choices that are determined by a variety of factors. Hence, the specific context and the specific problem studied might require different measures. Two interesting recent studies, both implemented in India, construct innovative measures that allow the identification of interesting parameters. Giannola (2023) studies how parents with a limited amount of resources allocate parental investment among different children whose initial perceived ability might be heterogeneous. Andrew and Adams (2022) model Indian parents' behavior on the marriage market for their daughters. We elaborate on these studies in Supplemental Appendix A.

The studies of choices and resource allocation within the household that we have mentioned are examples where the design and collection of new measures, guided by theory and validated in combination with choice data, can provide important insights.

**The determinants of parenting behavior.** Recently, the early years of human development have received a considerable amount of attention, partly following the work of Heckman and collaborators, some of which we mention in Supplemental Appendix A. It is now widely accepted that parental behavior plays a key role in determining early child development.

Standard models of child development and parenting behavior imply that the latter depends on resources, preferences, and, crucially, perceptions of the process of child development. Although a commonly used assumption is that parents fully understand the process of child development and the role their own behavior plays in it, it is not completely obvious that this is the case.

An alternative narrative allows parents to have *biased beliefs* about the process of child development. Lareau (2011) and Putnam (2015) discuss qualitative evidence indicating that disadvantaged parents invest much less time than better-off ones, maybe because they do not fully comprehend the process of child development and the importance of stimulation.[11] From a quantitative point of view, distinguishing between features of preferences and of the beliefs process requires obtaining information on these variables. Eliciting information about parental beliefs on the process of child development is the strategy taken in a number of recent papers.

Cunha, Elo, and Culhane (2022) have developed a methodology to elicit such beliefs and used it in a sample of disadvantaged mothers in Philadelphia. Attanasio, Cunha, and Jervis (2019) further developed that approach in a survey evaluating the impact of a parental intervention in Colombia. These papers, and others mentioned in Supplemental Appendix A, show that parental beliefs can be measured and that these beliefs seem to be informative of actual parental behavior. Cunha et al. (2022) and Attanasio et al. (2019) put more structure on the beliefs data and use

[11] Bellue (2023) develops a general equilibrium model of location where distorted beliefs about the process of child development arise in some neighborhoods as an equilibrium in which there is no learning towards an unbiased belief. An analogous model is presented by Dasgupta and Saha (2022).

them to estimate a *subjective* production function of human capital. These estimates can then be compared to estimates derived from data on actual child development and parental investment (along other factors). Such a comparison allows researchers to establish the extent to which parents under- or overestimate the productivity of parental investment. An important point to notice in the present context is that, while estimating the *objective* production function requires accounting for the endogeneity of parental investment (since it is determined by parental *choices*), with data on subjective beliefs endogeneity issues are not a concern, as these estimates are derived from counterfactual data which have embedded in them the necessary variation in individual inputs.

**Education choices.** Investment behavior in a variety of contexts — ranging from the allocation of individual portfolios to different assets to the education choices of high school and university students and to parenting practices — is likely to be driven by several factors, including, crucially, the information available to investors about the nature of returns on different investment opportunities. It is well established that individuals may not have accurate information about the expected value and riskiness of these opportunities; therefore, obtaining information on this key factor in individual choices is useful.

An area where many researchers have collected and used information on perceived returns is that of returns to education choices. Two of the first papers in this context are Dominitz and Manski (1996, 1997). In a more recent paper, Arcidiacono et al. (2020) leverage subjective expectations data on future earnings in a model of occupational choice. Giustinelli (2023) provides a recent review of this literature.

Several studies, some of which we mention in Supplemental Appendix A, have considered the role of subjective expectations in education choices. An important thing to stress here is the flexibility in modeling individual choices when a wider set of measures is available. At the same time, as we already mentioned, the availability of subjective belief data does not necessarily solve the endogeneity problems. This point is clear in Wiswall and Zafar (2015). The authors conduct an RCT of an information intervention which introduces exogenous variation into the belief about future earnings held by students choosing a college major.

**Learning.** Some of the examples we have discussed above, such as models of parental behavior, and school choices, but also earnings and employment dynamics, once they depart from the assumption of rational expectations suggest the need to model how individual perceptions and expectations change in response to signals. Learning has received considerable attention in economic theory (see, for instance, Evans and Honkapohja, 2001; Fudenberg and Levine, 1998). From an empirical point of view, a few papers have tried to model learning, using choice data and often assuming rational expectations and Bayesian updating, starting with Miller (1984), and, more recently, Arcidiacono, Aucejo, Maurel, and Ransom (2025). Other contributions (Stinebrickner and Stinebrickner, 2014a,b; Wiswall and Zafar, 2015) used elicited beliefs to model how individuals update subjective beliefs. Much of this work is about students' perception about their own ability and how certain signals affect them. However, there is little additional empirical work on how individual beliefs and expectations change in response to new information and signals.[12]

**Firm behavior: managerial ability and subjective expectations.** The standard economic model of production assumes that firms use physical and human capital (which might come in different dimensions) to produce value added from a set of inputs, such as raw materials and intermediate goods. More recently, managerial skills have been considered an additional factor in production. The seminal work of Bloom and Van Reenen (2007) introduced innovative measures of management skills which have been collected in a wide range of surveys carried out in many countries. These contributions are further discussed in Bloom, Lemos, Sadun, Scur, and Van Reenen (2014) and Scur, Sadun, Van Reenen, Lemos, and Bloom (2021), as well as other papers, some of which we cite in Supplemental Appendix A.

Firms' choices, including investment decisions, are obviously influenced by their expectations of market conditions and economic activity, and they are also likely influenced by uncertainty and risk. Measures of subjective perceptions of risk could therefore be useful when empirically modeling these choices, as shown in several papers, such as Bloom, Bond, and Van Reenen (2007) and Mohnen, Hall, Mairesse, and Mohnen (2008). Recently, Henzel (2021) investigated the role of subjective uncertainty and expectations in shaping firm behavior, with a particular focus on investment and employment decisions during the COVID-19 crisis, while Fiori and Scoccianti (2023) used two decades of Italian survey data on business managers' expectations to empirically characterize subjective firm-level uncertainty and quantify its economic effects.

In a new paper, Norris Keiller, de Paula, and Van Reenen (2024) use data on subjective expectations at the firm level,[13] to identify the structural parameters of a production function and address the important issue that managers choose inputs based on information not available to the econometrician. Although the preexisting literature, such as Olley and Pakes (1996), addresses this issue with assumptions on the timing of different inputs, this study takes advantage of subjective expectations data to avoid relying on such restrictions and to identify dynamic models with

[12] An exception is work in experimental economics (see, for instance, Nagel (1995)). Some studies, such as Malmendier and Nagel (2011), characterize how some aspects of preferences are shaped by life experiences.

[13] They use Management and Expectations Survey (MES), a survey administered by the UK's statistical authority, whose structure is similar to the data used in Bloom and Van Reenen (2007). The Bank of England collects another UK survey on expectations of firms' CFOs, the Decision Maker Panel.

relatively short panels.[14]

**Analyzing deviations from benchmark models: rational expectations.** We have repeatedly stressed how novel measurements can test and quantify deviations from benchmark models. Our empirical application — using new data on social cohesion and income perceptions to assess alternative risk-sharing models — demonstrates this potential. The Rational Expectations (RE) literature offers a parallel case. As discussed in Lovell (1986), early debates questioned whether RE is testable at all, echoing our point that the viability of identification strategies depends critically on both the model's structure and the available data.

With the growing use of subjective expectations data, interest in testing the rational expectations hypothesis (RE) has rekindled. D'Haultfoeuille, Gaillac, and Maurel (2021) propose a new test based on the marginal distributions of outcomes and beliefs, finding support for RE among college graduates but significant deviations for others. Crossley, Gong, Stinebrickner, and Stinebrickner (2024) show that incorporating higher moments of belief distributions uncovers RE violations missed by mean-based tests, including downward-biased uncertainty among young college graduates in the Berea Panel Study.

### 2.3 Challenges

The studies cited above illustrate a broader effort to develop empirical measures that capture complex behavioral models. While such tools hold great promise, their construction poses significant challenges. The theoretical concepts they aim to quantify — though grounded in models of human behavior — are often difficult to define and measure consistently. As a result, careful design and rigorous validation are essential for their effective use.

Psychometrics offers a rich literature on validating measurement tools (see, e.g., Messick, 1995; DeVellis, 2016), from which economists can draw valuable lessons. However, economic applications raise additional concerns: validation must also assess how well measures align with theoretical constructs and relate to observed choices and decisions. In what follows, we highlight some of the most pressing challenges in this context.

**Measurement error and establishing the metric of latent factors.** New measures are likely to be affected by measurement and elicitation errors. Indeed, the very recognition that these measures might be an attempt to get estimates of unobservable *latent factors* recognizes explicitly the existence of measurement error, such as in the measurement systems considered in equation (3). In

[14] The argument is analogous to the one used by Arellano et al. (2025) and discussed in Section 2.1.

this respect, the use of latent factor models is becoming a common technique for synthesizing and aggregating available measures. Early contributions in economics on these issues include Goldberger and Duncan (1973) and Goldberger and Jöreskog (1975). Goldberger (1972) is particularly relevant because it combines the discussion of measurement error and the existence of complex causal links among several variables.

A common challenge in economics and other sciences is defining a meaningful scale for latent factors measured through observed indicators. Building on Cunha et al. (2010), Agostinelli and Wiswall (2025) examine this issue in the context of modeling child development, where ensuring comparability of measures across ages is crucial for estimating the dynamics of skill formation. This case is especially noteworthy because the standard psychometric approach — normalizing measures at each age — conflicts with the focus in economic models, which emphasize tracking skill growth over the life cycle. [15]

An additional difficulty faced by researchers in establishing a metric for certain variables of interest arises when all the available measures are discrete. In such a situation, a latent factor model such as Item Response Theory is often used within Structural Equation Modeling (see Bollen, 1989). However, standard algorithms impose strong functional form assumptions on the distribution of the unobserved latent factors (such as normality).

**Anchoring expectations questions.** As previously noted, substantial progress has been made in the elicitation and use of data on subjective expectations. It is now well established that researchers can obtain meaningful measures of individual expectations that capture the probability distributions respondents associate with uncertain future outcomes. Moreover, with careful piloting, it is possible to collect probabilistic information in a credible and reliable manner. Nonetheless, because the most effective protocols for gathering high-quality data are often context-dependent, designing questions about future income anchored to specific, interpretable values remains challenging.

In the case of expectations regarding future household income, one common approach—used, for instance, in the New York Fed Survey of Consumer Expectations—is to first ask respondents for a point estimate and then follow up with questions designed to elicit the cumulative distribution function (CDF) of income around that estimate. An alternative strategy is to anchor expectations to the respondent's current income level. A third approach involves asking for a range of values that the respondent considers to approximate the minimum and maximum possible future income. Each of these methods has its own strengths and limitations. Given these challenges—and

[15] The normalization of measures of child development to establish a metric across ages is also discussed in Attanasio, Bernal, Giannola, and Nores (2020). While Cunha et al. (2010) use adult wages as a natural metric, these more recent articles limit themselves to childhood outcomes. See also the discussion in Freyberger (2024).

recognizing that the optimal question design is likely to vary by context—it would be valuable to develop a standardized protocol for testing and validating such survey instruments.

**The design of counterfactual questions.** The use of stated preferences or intentions in hypothetical situations has, for a long time, been received with skepticism in economics (see for instance Tobin (1959)). While this has been changing and an increasing number of surveys contain this type of information, some of which we have mentioned above, the design and use of such questions is not without challenges. Even more than with other questions, extensive piloting of this type of questions is necessary. Furthermore, the type of counterfactual scenarios one wants to use should be informed by specific theoretical questions. For example, Arellano et al. (2025) make extensive use of data on the probability distribution of potential offers that an individual might receive, which, within the model they are considering, allows them to identify key parameters avoiding some of the identification issues of discrete choice models.

At the same time, to identify the parameters of interest, such data, while useful, needs to be complemented by a set of assumptions, possibly different from those used when using choice data. The type of variation provided by the hypothetical choice data might be discrete in nature, implying that in some contexts identification might require distributional and/or functional form restrictions.

**Learning and the strength of priors.** As noted above, empirical research on learning under uncertainty remains limited. While it is intuitive that agents process signals to improve their forecasts, direct evidence on learning mechanisms is scarce. Even studies such as Cunha et al. (2022) and Attanasio et al. (2019), which document distortions in beliefs about specific processes, do not explicitly model how beliefs update in response to new information. We notice that from an empirical point of view, the data requirements can be demanding: one would need longitudinal data on beliefs and/or expectations and information about potential "signals" individuals observe.

Standard models of learning—such as Bayesian learning—imply that posterior beliefs depend on both the signals received and the strength of prior beliefs. The influence of priors hinges on the confidence with which they are held, yet surveys rarely collect this kind of information, even when they elicit expectations. An exception is Giustinelli et al. (2019), which distinguishes between precise and imprecise probabilities and accounts for rounding behavior.[16]

[16]Capturing the degree of conviction with which beliefs are held would be useful for modeling learning and evaluating information interventions. However, challenges arise when respondents report intervals rather than point estimates. In these cases, key parameters—like the mean—are only partially identified, with the sharp identified set given by the Aumann expectation of the reported intervals (see Aumann, 1965). This issue is particularly relevant given the widespread occurrence of such "imprecise" beliefs. We thank Arnaud Maurel for highlighting this point.

**Belief heterogeneity.** While heterogeneity in preferences has long been acknowledged in economics—see, for example, the aggregation results in Harsanyi (1955)—less attention has been paid to heterogeneity in beliefs about future uncertainty. However, as pointed out in Gilboa, Samet, and Schmeidler (2004), heterogeneous beliefs about the probability distribution of future outcomes can be very important for the characterization of equilibria. Empirical measures of belief heterogeneity could provide valuable insights and give empirical traction to these theoretical frameworks.

This is particularly relevant in contexts such as intrahousehold resource allocation, including parental investments in children. Extending collective models (e.g., Chiappori (1988)) to incorporate heterogeneous beliefs is a promising avenue. Models could also incorporate belief distortions of the kind studied in Bellue (2023), especially when panel data allow for tracking how individuals learn and how beliefs interact within households or risk-sharing groups over time.

**Measuring the quality of information.** Information frictions are often cited as a key reason why risk sharing appears inefficient in many settings. These frictions may arise from moral hazard, adverse selection, or simply from the lack of publicly observable information about idiosyncratic income. Although theoretical models have explored equilibria under imperfect information, empirical evidence on the consequences of such frictions remains limited. A notable exception is Attanasio and Krutikova (2020), discussed further below. Developing and validating new empirical measures of information quality would be a valuable contribution.

## 3 Imperfect risk sharing and measures to characterize it

In this section, we provide an example in which the existence of new measures can shed light on economic models that might be richer than those relying on strong assumptions and that might be rejected by empirical data. We build on the extensive literature on risk sharing in rural economies. Starting with the groundbreaking work of Townsend (1994), a large literature has considered the empirical implications of a model in which idiosyncratic risk within a predefined group is perfectly insured.

One appealing feature of Townsend's approach is that it characterizes the empirical implications of consumption allocations under perfect risk sharing within a group, without specifying the market mechanisms that implement those allocations or defining the precise nature of the risk-sharing group. By comparing actual consumption patterns to this theoretical benchmark, we can assess the extent of risk sharing and quantify deviations from the first-best allocation.

The basic idea in Townsend's approach is that, given a risk-sharing group, after controlling for group-specific time effects, idiosyncratic shocks to income are not reflected in changes in the marginal utility of consumption, which can be approximated by log consumption. We interpret the strength with which this implication of the model is rejected, that is the extent to which idiosyncratic shocks to income are reflected into consumption, as a measure of how much consumption allocations are distant from perfect risk sharing.

The next step is to consider theoretical constructs that might be helpful in understanding deviations from perfect risk sharing. We first hypothesize that to implement perfect risk sharing, a certain amount of *social cohesion* is necessary. Second, we consider the fact that one of the key assumptions needed for perfect risk sharing to be decentralizable is perfect information. We therefore consider the relation between measures of the quality of information within a risk-sharing group and deviations from first-best. Finally, we consider the lack of enforceability of contracts as a possible friction preventing perfect risk sharing and notice that, in these models, the extent of risk sharing can be related to properties of the income process.

The Townsend approach can be derived from a social planner problem, maximizing the expected utility of the risk-sharing group members, with a fixed Pareto weight given to each of them. As shown in detail in Supplemental Appendix B, assuming power utility, and taking logs of the marginal utility of consumption, one obtains the following expression for the log level of individual consumption in risk-sharing group $v$, $\ln c^v_{i,t}$

$$\ln c^v_{i,t} = \frac{1}{\gamma}(t \ln \beta - \ln \theta^v_t + \ln \lambda^v_i) \quad \forall i, t. \tag{5}$$

where $\lambda^v_i$ is the Pareto weight given to individual $i$, which does not change with time, $\beta$ the discount factor and $\theta^v_t$ the Lagrange multiplier associated with the group resource constraint at time $t$. Taking the first difference of equation (5), and taking into account that the Pareto weight is constant over time, we get:

$$\Delta \ln c^v_{i,t} = \frac{1}{\gamma}(\ln \beta - \Delta \ln \theta^v_t) \qquad \forall i, t. \tag{6}$$

We notice that the right-hand side of equation (6) does not depend on the individual subscript $i$, so that perfect risk sharing implies that consumption growth depends only on changes in the *aggregate multiplier* $\Delta \ln \theta^v_t$ associated with the resource constraints for group $v$. The implications of equations (5) and (6) have been extensively tested in the literature by the following regressions:

$$\ln c^v_{i,t} = \pi_{0l,i} + \delta^{vl}_t + \pi_{1l} \ln y^v_{i,t} + \varepsilon^{vl}_{i,t} \tag{7}$$

$$\Delta \ln c^v_{i,t} = \pi_{0d} + \delta^{vd}_t + \pi_{1d} \Delta \ln y^v_{i,t} + \varepsilon^{vd}_{i,t} \tag{8}$$

where $\delta^{vl}_t$ and $\delta^{vd}_t$ are group and time dummies interacted, $\pi_{0l,i}$ are individual fixed effects and

$\ln y^v_{i,t}$ is idiosyncratic income levels. Perfect risk sharing implies $\pi_{1l} = \pi_{1d} = 0$. These coefficients measure how idiosyncratic income shocks are reflected into consumption changes, given group-level income shocks. The size of the coefficients, therefore, can be interpreted as the extent of the deviation from first-best allocations.

Of course, equations (7) and (8) can be enriched to take into account a variety of other factors. First, one could consider other observable variables that appear related to changes in consumption and that could be explained by the theory as determinants of the marginal utility of consumption. Second, variables other than changes in (log) income could be considered as proxying for idiosyncratic shocks. Third, as it has been done in the literature, one could consider heterogeneity in preferences and attitudes towards risk, such as Schulhofer-Wohl (2011). Finally, and important for our discussion, one could consider models of specific frictions that could mediate the observed deviations from perfect risk sharing. In what follows, we consider measures of social cohesion, of the quality of information, and of the properties of income processes.

**Failures of perfect risk-sharing tests.** There are many reasons why perfect risk-sharing tests, implemented through equations (7) and (8), can reject the null of perfect risk sharing. One possibility is that these rejections arise from misspecification of the underlying model, such as the assumption of power utility or homogeneity in preferences. Alternatively, failures of perfect risk sharing could stem from specific frictions that prevent first-best allocations. In Supplemental Appendix B.1, we discuss in more detail these possibilities and part of the literature that has developed around them. Here we limit ourselves to listing those frictions for which we consider some novel measures as possible mediators to explain the empirical failures of perfect risk sharing.

We first discuss the idea that risk sharing presumes the existence of social cohesion and trust and refer to some of the studies that have considered measures of social capital. The next friction we consider is that of imperfect information. While this type of imperfection has received considerable attention in the theoretical literature, the empirical evidence on this — which we discuss in Supplemental Appendix B.1 — is limited, in part because of the scarce availability of measurements of the *quality of information.* Finally, we consider models where insurance and risk-sharing contracts are not perfectly enforceable. We note that in such models, the degree of risk sharing depends on the properties of the income process individuals in a risk-sharing group face.

# 4 New measures for risk sharing models: some examples

In what follows, we use micro data to test the hypothesis of perfect risk sharing and then introduce three novel measures that may add interesting insights to characterize deviations from the null of complete insurance markets. In particular, exploiting a very rich dataset collected several years ago in Colombia, we use measures on social cohesion, information quality, and properties of income processes that could be relevant for the enforcement of contracts.

The data we use come from a survey conducted in 122 Colombian villages across 26 of the country's 34 departments, as part of the evaluation of the first phase of the conditional cash transfer program *Familias en Acción*. In each village, approximately 100 eligible households were sampled. The baseline survey, carried out in 2002, included interviews with 11,462 households. Two follow-up surveys were conducted in 2003 and 2005–2006, re-interviewing 10,743 and 9,463 of the baseline households, respectively. Our main sample is an unbalanced panel comprising 11,914 unique households. Further details on the data are provided in Supplemental Appendix C.

### 4.1 Subjective expectations of income processes

Households choose consumption and savings plans, and might enter risk-sharing agreements by weighing their own *perceptions* about their future income. Subjective expectations data provide households' own assessments of future income and are directly informative about the perceived features of the income processes they face. Following Arellano, Attanasio, Crossman, and Sancibrián (2024), we take advantage of the presence of both income expectations and realizations in the two follow-up rounds in the Colombian survey data to obtain village-level summaries of subjective income risk and persistence.

Respondents (indexed by $i$) in survey round $t$ were asked to provide a range of variation for their future income $y^v_{i,t+1}$, and then report the probability $p_{jit}$ that $y^v_{i,t+1}$ will not exceed each of three equally-spaced points $r_{jit}$ for $j = \{1, 2, 3\}$ within the elicited range of variation. With a log-linear specification for the individual income process and some assumptions on the elicitation process, Arellano et al. (2024) relate the logit-transformed subjective probabilities $\ell_{jit} = \text{logit}(p_{jit})$ to the parameters that characterize the income process, obtaining the following linear specification:[17].

$$\ell_{jit} = \beta_0 r_{jit} + \beta_1 y_{i,t} + \mu_i + u_{jit}, \tag{9}$$

where $\mu_i$ are fixed effects and $u_{jit}$ are elicitation errors. Arellano et al. (2024) show that the parameters of equation (9) map directly to the "structural" parameters $\{\rho, \sigma\}$ of the first-order autoregressive process in equation (4) with logistic innovations, with $\beta_0 = 1/\sigma$ and $\rho = -\beta_1/\beta_0$.

We estimate equation (9) village-by-village and obtain village-specific measures of subjective risk and persistence. Table 1 reports summary statistics of these estimates. As we drop villages

[17] Arellano et al. (2024) consider also more flexible, non-linear specifications for the income process

with fewer than 20 households, we obtain estimates of the parameters for 86 of the 122 villages.[18]

| | Obs. | Mean | Std. dev. | 10th perc. | 50th perc. | 90th perc. |
|---|---|---|---|---|---|---|
| $\sigma_v$ | 86 | 0.508 | 0.222 | 0.239 | 0.505 | 0.796 |
| $\rho_v$ | 86 | 0.553 | 0.141 | 0.394 | 0.530 | 0.743 |

TABLE 1. Summary statistics for $\rho_v$ and $\sigma_v$ for 86 villages.

For the 86 villages, $\rho$ is estimated to av erage 0.508, a mean that coincides with that reported in Arellano et al. (2024, Table D.3) using the pooled full sample. However, the estimated persistence parameter exhibits considerable variability across villages: the 10th percentile is as low as 0.24, while the 90th percentile is 0.8. Turning to estimates of subjective risk, the average standard deviation reported in Table 1 is high, at 0.55. Again, there is considerable variability across villages, with the 10th percentile at 0.39, while the 90th percentile is at 0.74.

### 4.2 Experimental data

The main survey contains data on consumption and current and expected income. In addition, the data set also contains a variety of measures that can be related to social cohesion and information quality, obtained in a lab-in-the-field experiment implemented in 70 of the 122 Colombian villages. Within each of the 70 villages, a random sample of 60 households was invited to send a household member to participate in the laboratory experiments. On average, 40 people per village showed up, for a total of 2,512 participants. Participants in the experiment were asked to play first a public goods game and then a risk-sharing game. Details on the experimental design are available in Attanasio, Barr, Cardenas, Genicot, and Meghir (2012a); here we briefly describe the two games.

**Public goods game.** In the public goods game, each participant receives a token worth 2,000 Colombian pesos and decides, privately, whether to keep it or contribute it to a public account. Contributions to the public account are doubled in value to 4,000 pesos. At the end of the game, the public account is evenly divided among the participants, who also retain any tokens they choose to keep. The social optimum is for everyone to put the voucher in the public account, while the Nash equilibrium is for everybody to keep the voucher for themselves. The game is played in two rounds, with participants allowed to discuss strategies between rounds, though outcomes are only revealed after the second round.

[18] We report robustness checks in Supplemental Appendix D.

| | Obs. | Mean | Std. dev. | 10th perc. | 50th perc. | 90th perc. |
|---|---|---|---|---|---|---|
| $VCM_v(1)$ | 2166 | 0.367 | 0.482 | 0 | 0 | 1 |
| $VCM_v(2)$ | 2166 | 0.441 | 0.497 | 0 | 0 | 1 |
| $IC_i$ | 2374 | 2.584 | 2.324 | 0.500 | 2 | 5.750 |
| $OC_i$ | 2374 | 2.786 | 2.096 | 0.500 | 2.250 | 5.750 |
| $\overline{size}_v$ | 2374 | 5.205 | 5.791 | 1 | 4 | 9 |

TABLE 2. Summary statistics for the collected measures from the experiment.

**Risk-sharing game.** In the risk-sharing game, participants first individually play a game where they choose one of six lotteries that are different in expected value and risk: the first lottery delivers a fixed amount with certainty, while the five lotteries deliver higher expected values but with some risk. After that, participants are given the possibility to form risk-sharing groups, which are registered. After playing individually, the results of the individual lotteries are shared among the group members, although individuals are given the option of privately reneging on their group membership.

With the data collected in the laboratory experiment, we create measures for social cohesion and information quality, based on the hypothesis that these variables could sustain risk sharing. These measures were calculated for each of the 70 villages in which the experiments were performed.[19]

**Measures of social cohesion.** For each village, we calculate three measures of social cohesion using data from the public goods game: (i) $VCM_v(1)$, the proportion of individuals in village $v$ who contributed to the public account in the first round; (ii) $VCM_v(2)$, the proportion of individuals in village $v$ who contributed to the public account in the second round; (iii) $\overline{VCM}_v$, the average of $VCM_v(1)$ and $VCM_v(2)$.

Using data from the risk-sharing game, in each village $v$, we calculate an additional measure of social cohesion: the average size of the risk-sharing group ($\overline{size}_v$). If risk preferences are homogeneous, the optimal strategy would be to form one large risk-sharing group. In practice, many small risk-sharing groups are formed, as discussed in Attanasio et al. (2012a).

**Measures of information quality.** In the lab-in-the-field experiment, some information about the existence of links among participants was collected between the public good game and the risk-

[19]Information on the choices made in the first part of the risk-sharing game could be used to construct individual-level measures of risk aversion, which might allow us to explore preference heterogeneity in risk sharing. Preliminary results do not suggest a large role for differential risk-aversion in explaining the rejection of perfect risk sharing in our context. For this reason, we do not pursue this avenue.

sharing game. Interviewers asked each participant if they knew each of the other participants and the nature of the relationship (acquaintanceship, friendship, family relative). With this information, we construct adjacency matrices[20] and, following Attanasio and Krutikova (2020), we construct measures of the quality of information in each village. Given the adjacency matrix for a village, we compute indexes of inward ($IC_i$) and outward degree centrality ($OC_i$), the former measuring how well-known individual $i$ is by the rest of the community, and the latter measuring how close their relationship is with the rest of the community. The adjacency matrices are not necessarily symmetric; hence the two centrality measures may be different.

**Descriptive statistics.** In Table 2, we report descriptive statistics of the measures of social cohesion and information quality in the 70 villages where the experimental games were performed. The average share of people contributing to the public goods game increases from the first to the second round by around 7 percentage points (from 37% in the first round to 44% in the second round). Moving to the two measures of network centrality, note that their means are not very different, suggesting that the adjacency matrices might not be far from symmetry. Compared to the inward centrality measure, outward centrality distribution is more concentrated around its mean. On average, risk-sharing groups are formed by 5 members, considerably lower than one would expect with full risk sharing under similar risk aversion. Finally, note that there is substantial heterogeneity across villages in average group size.

# 5 Empirical evidence on richer models of risk sharing

In this section, we report the empirical results using the survey and experimental data from Colombia described above. We begin by estimating equations (7) and (8) to study empirical deviations from perfect risk sharing. Table 3 reports estimates of $\pi_{1l}$ and $\pi_{1d}$, measuring the effect of idiosyncratic income shocks after controlling for village and time effects in the difference specification and also controlling for household fixed effects in the level specification.

The regression results point to statistically significant coefficients in both levels and first differences, indicating a strong rejection of perfect risk sharing. In both specifications, a 1% increase in income is significantly associated with a 0.16% increase in consumption.

**Risk-sharing and social cohesion.** To improve our understanding of risk-sharing arrangements, we test the hypothesis that groups with better cohesion are closer to perfect risk-sharing than

[20] The $i, j$ element of the adjacency matrix is 1 if $i$ and $j$ are family, 0.5 if they are friends, and 0.25 if they are acquaintances.

| | (1) $\ln c_{it}^{v}$ | (2) $\Delta\ln c_{it}^{v}$ |
|---|---|---|
| $\ln y_{it}^{v}$ | 0.160*** | |
| | (0.008) | |
| $\Delta\ln y_{it}^{v}$ | | 0.157*** |
| | | (0.008) |
| Household FE | Yes | No |
| Village × time FE | Yes | Yes |
| Adj. $R^2$ | 0.565 | 0.145 |
| Observations | 25379 | 14777 |

TABLE 3. Baseline Townsend (1994) tests. Time refers to survey rounds. Standard errors clustered at the village level. *10%, ** 5%, ***1%.

groups with lower cohesion. To measure social cohesion, we use both the results from the public goods games, that is, variables $VCM_v(1)$, $VCM_v(2)$ and $\overline{VCM}_v(1)$ — as well as the average size of the risk-sharing groups formed.

Table 4 reports the estimates of coefficients of a modified version of equation (8), where the first differences in the (log of) household income interacts with the measures of social cohesion. The coefficients on the various interaction terms are not statistically different from zero, indicating that our measures of social cohesion do not seem to be related to deviations from perfect risk sharing. We obtain similar results for the specifications in levels (see Supplemental Appendix D). One possibility is that these measures do not capture the variables that are relevant to the implementation of risk sharing. Another possibility is that our social cohesion variables are measured with error.[21]

**Risk sharing and imperfect information.** The second hypothesis we consider is the possibility that the quality of information influences the observed levels of risk-sharing. Using the measures of inward and outward centrality collected in the lab-in-the-field experiment, we estimate a modified version of equations (7) and (8) by interacting these measures with household income.

Columns (1) and (2) in Table 5 report the results. Our measures of information quality seem to be systematically related to the extent of deviations from perfect risk-sharing. Households with

[21]The availability of repeated measurements of the same underlying quantity might help solve such measurement problems, as we discussed in Section 2.1. This setup provides one tentative application: we have two separate measurements of social cohesion, one from each round of the public goods game. Provided that both are informative enough and that measurement errors across game rounds are uncorrelated, we can use one measure as an instrument for the other. In this case, instrumental-variable estimates are similar to those reported in Table 4, suggesting that (classical) measurement errors might not be driving the results.

| | (1) $\Delta \ln c_{it}^v$ | (2) $\Delta \ln c_{it}^v$ | (3) $\Delta \ln c_{it}^v$ | (4) $\Delta \ln c_{it}^v$ |
|---|---|---|---|---|
| $\Delta \ln y_{it}^v$ | 0.151*** | 0.148*** | 0.148*** | 0.154*** |
| | (0.018) | (0.018) | (0.020) | (0.012) |
| $\text{VCM}_v(1) \times \Delta \ln y_{it}^v$ | 0.010 | | | |
| | (0.039) | | | |
| $\text{VCM}_v(2) \times \Delta \ln y_{it}^v$ | | 0.016 | | |
| | | (0.028) | | |
| $\overline{VCM}_v \times \Delta \ln y_{it}^v$ | | | 0.018 | |
| | | | (0.040) | |
| $\overline{size}_v \times \Delta \ln y_{it}^v$ | | | | 0.000 |
| | | | | (0.001) |
| Household FE | No | No | No | No |
| Village × time FE | Yes | Yes | Yes | Yes |
| Adj. $R^2$ | 0.147 | 0.147 | 0.147 | 0.147 |
| Observations | 10419 | 10419 | 10419 | 10419 |

TABLE 4. Townsend (1994) tests augmented with interactions with measures of social cohesion. In Columns (1) and (2) social cohesion is measured using the share of individuals in each village contributing to the first and second round of the public goods game respectively, while in Column (3) we use the average of the two measures. Column (4) includes the average risk-sharing group size as a measure for social cohesion. Time refers to survey rounds. Standard errors clustered at the village level. *10%, ** 5%, ***1%.

high inward or outward centrality are less affected by idiosyncratic income shocks. If one interprets the degree centrality measures we use as a proxy for information quality in the risk-sharing groups considered (the participants to the lab-in-the-field experiment in each village), these results imply that information imperfections might be behind the failure of perfect risk sharing.

**Imperfect enforceability: risk sharing and income processes.** Next, we explore how deviations from perfect risk-sharing vary with the (perceived) persistence and dispersion of the income processes households face across villages.

In Column (3) of Table 5, we augment the baseline Townsend equations with interactions with village-level perceived risk and persistence $\{\sigma_v, \rho_v\}$, identified from subjective expectations data.[22]

The estimates suggest a role for the characteristics of (perceived) income processes to mediate the deviations from perfect risk-sharing that we documented in the villages included in our sample.

[22]We report results in levels and robustness checks in Supplemental Appendix D. As discussed in Section 4, $\{\sigma_v, \rho_v\}$ are estimated in a first-step and non-negligible estimation noise might remain, which could lead to attenuation bias in our estimates. We exclude villages with fewer than 20 unique households from the analysis. Table D.5 in Supplemental Appendix D reports robustness checks dropping only villages with fewer than 10 unique households and reports very similar results.

| | (1) $\Delta \ln c_{it}^{v}$ | (2) $\Delta \ln c_{it}^{v}$ | (3) $\Delta \ln c_{it}^{v}$ |
|---|---|---|---|
| $\Delta \ln y_{it}^{v}$ | 0.208*** | 0.215*** | 0.223*** |
| | (0.022) | (0.025) | (0.035) |
| $\text{IC}_i \times \Delta \ln y_{it}^{v}$ | -0.014** | | |
| | (0.007) | | |
| $\text{OC}_i \times \Delta \ln y_{it}^{v}$ | | -0.016** | |
| | | (0.007) | |
| $\sigma_v \times \Delta \ln y_{it}^{v}$ | | | -0.200*** |
| | | | (0.064) |
| $\rho_v \times \Delta \ln y_{it}^{v}$ | | | 0.100*** |
| | | | (0.033) |
| Household FE | No | No | No |
| Village × time FE | Yes | Yes | Yes |
| Adj. $R^2$ | 0.160 | 0.160 | 0.147 |
| Observations | 3467 | 3467 | 13428 |

TABLE 5. Townsend (1994) tests augmented with interactions with centrality measures (Columns 1 and 2) and village-level measures of subjective risk and persistence (Column 3). $IC_i$ is calculated as the weighted number of individuals that recognized individual $i$ as a family/friend/acquaintance. $OC_i$ is calculated as the weighted number of individuals recognized by individual $i$ as a family/friend/acquaintance. Time refers to survey rounds. Standard errors clustered at the village level. *10%, ** 5%, ***1%.

In particular, these deviations are smaller in villages where income processes are perceived to be less persistent or more volatile. These results are consistent with the intuition from models with imperfectly enforceable contracts that we discuss in Supplemental Appendix B.1.

## 6 Conclusion

In this paper, we discuss the value of using new and innovative measures in empirical work in economics. The profession, with some exceptions, has been skeptical about using data on stated preferences, answers to questions about choices in hypothetical situations, attitudes, and subjective beliefs and expectations. An important literature on the exclusive use of choice data and what can be learned from such data with a limited number of assumptions has developed in foundational papers by Afriat (1967) and Varian (1982) and subsequent contributions, some of which are reviewed in Dziewulski, Lanier, and Quah (2024) and Chambers and Echenique (2016). However, this approach can be restrictive in terms of the models that can be studied and the assumptions needed to identify the parameters of interest.

In the first part of the paper, we review a number of contexts where innovative measures facilitate the identification of structural parameters under a considerably weaker set of assumptions than would be required in their absence, and others where rich models of human behavior can be characterized empirically. In the second part, we illustrate some of our points within a specific context. We start from the rejection of a restrictive model of insurance: one assuming perfect risk sharing within a certain group of idiosyncratic shocks. We then show how additional measures can capture economic dimensions that might hinder perfect risk sharing, and we investigate whether these factors are empirically relevant using survey data from Colombia. We show that while some of these factors do not appear to explain deviations from perfect risk sharing, others are systematically related to such deviations and the observed correlations are consistent with specific models of imperfect risk sharing, such as those with imperfectly enforceable contracts.

Fortunately, the profession's stance on these issues has been changing, and new measures are often collected and developed. Data on subjective expectations have become common, and their collection has improved. New measures are now trying to capture different aspects, such as the strength with which certain views are held. In addition, researchers are now using data derived from stated preferences and hypothetical choices. New methods are also being developed to use qualitative data — such as answers to open-ended questions — in quantitative studies.

We conclude with an important caveat. Although innovative and original measures could be extremely useful, it should be recognized that some theoretical constructs are inherently difficult to measure. When designing and eventually implementing new measures, extensive validation is important. Much research effort is needed, using insights from many disciplines, including psychology, survey methodology, marketing, and statistics. New measures may be subject to various biases and distortions. To ensure their credibility, they should be thoroughly validated.

## Data availability

Code replicating the tables in this article can be found in Attanasio, Sancibrián, and Ambrosio (2025) in the Harvard Dataverse, `https://doi.org/10.7910/DVN/SQ4XNB`.

## References

ABOWD, J. M. AND D. CARD (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445.

AFRIAT, S. N. (1967): "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 8, 67–77.

Agostinelli, F. and M. Wiswall (2025): "Estimating the Technology of Children's Skill Formation," *Journal of Political Economy*, 133, 631–639.

Alesina, A. and E. La Ferrara (2002): "Who trusts others?" *Journal of Public Economics*, 85, 207–234.

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

Almås, I., A. Cappelen, E. Sørensen, and B. Tungodden (2020a): "Fairness across the world: preferences and beliefs," Tech. rep., NHH working paper.

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

Almås, I., O. Attanasio, and P. Jervis (2024): "Presidential Address: Economics and Measurement: New Measures to Model Decision Making," *Econometrica*, 92, 947–978.

Altonji, J. G., A. A. Smith, and I. Vidangos (2013): "Modeling earnings dynamics," *Econometrica*, 81, 1395–1454.

Andrew, A. and A. Adams (2022): "Revealed beliefs and the marriage market return to education," Tech. rep., IFS Working Papers, No. 22/48.

Arcidiacono, P., E. Aucejo, A. Maurel, and T. Ransom (2025): "College Attrition and the Dynamics of Information Revelation," *Journal of Political Economy*, 133, 53–110.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020): "Ex Ante Returns and Occupational Choice," *Journal of Political Economy*, 128, 4475–4522.

Arellano, M., O. Attanasio, M. Borella, M. De Nardi, and G. Paz-Pardo (2025): "Subjective Earnings and Employment Dynamics," Tech. rep., In progress.

Arellano, M., O. Attanasio, S. Crossman, and V. Sancibrián (2024): "Estimating Flexible Income Processes from Subjective Expectations Data: Evidence from India and Colombia," Working Paper 32922.

Arellano, M., R. Blundell, and S. Bonhomme (2017): "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework," *Econometrica*, 85, 693–734.

Arellano, M. and S. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277–297.

Ashraf, N. (2009): "Spousal Control and Intra-household Decision Making: An Experimental Study in the Philippines," *American Economic Review*, 99, 1245–1277.

Athey, S. and G. Imbens (2017): "The Econometrics of Randomized Experiments," in *Handbook of Economic Field Experiments*, ed. by A. Banerjee and E. Duflo, Elsevier, vol. 1, 73–140.

Attanasio, O., A. Barr, J. C. Cardenas, G. Genicot, and C. Meghir (2012a): "Risk Pooling, Risk Preferences, and Social Networks," *American Economic Journal: Applied Economics*, 4, 134–67.

Attanasio, O., R. Bernal, M. Giannola, and M. Nores (2020): "Child development in the early years," Tech. rep., NBER, WP No 27812.

Attanasio, O., F. Cunha, and P. Jervis (2019): "Subjective Parental Beliefs: Their Measurement and Role," Working Paper 26516, National Bureau of Economic Research.

Attanasio, O. and S. Krutikova (2020): "JEEA-FBBVA LECTURE 2019: Consumption Insurance in Networks with Asymmetric Information: Evidence from Tanzania," *Journal of the European Economic Association*, 18, 1589–1618.

Attanasio, O., V. Sancibrián, and F. Ambrosio (2025): "Replication Data for: New Measures for Richer Theories: Some Thoughts and an Example." Harvard Dataverse https://doi.org/10.7910/DVN/SQ4XNB.

Attanasio, O. P., C. Meghir, and A. Santiago (2012b): "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA," *Review of Economic Studies*, 79, 37–66.

Aumann, R. J. (1965): "Integrals of Set-Valued Functions," *Journal of Mathematical Analysis and Applications*, 12, 1–12.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997): "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics*, 112, 537–579.

Bellue, S. (2023): "Residential and Social Mobility: A Quantitative Analysis of Parental Decisions with Social Learning," Working paper.

Bloom, N., S. Bond, and J. Van Reenen (2007): "The Economic Effects of Firm-Level Uncertainty: Evidence Using Subjective Expectations," NBER Working Paper 13327, National Bureau of Economic Research.

Bloom, N., R. Lemos, R. Sadun, D. Scur, and J. Van Reenen (2014): "The New Empirical Economics of Management," *Journal of the European Economic Association*, 12, 835–876.

Bloom, N. and J. Van Reenen (2007): "Measuring and Explaining Management Practices Across Firms and Countries," *The Quarterly Journal of Economics*, 122, 1351–1408.

Bollen, K. A. (1989): *Structural Equations with Latent Variables*, New York: Wiley.

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, 503–528.

Briggs, J. S., A. Caplin, S. Leth-Petersen, and C. Tonetti (2024): "Identification of Marginal Treatment Effects using Subjective Expectations," Working Paper 32309, National Bureau of Economic Research.

Browning, M. and T. Crossley (2009): "Are Two Cheap, Noisy Measures Better Than One Expensive, Accurate One?" *American Economic Review: Papers & Proceedings*, 99, 99–103.

Chambers, C. P. and F. Echenique (2016): *Revealed Preference Theory*, Cambridge University Press.

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

Chiappori, P.-A. (1992): "Collective Labor Supply and Welfare," *Journal of Political Economy*, 100, 437–467.

Crossley, T. F., Y. Gong, R. Stinebrickner, and T. Stinebrickner (2024): "Examining Income Expectations in the College and Early Post-College Periods: New Distributional Tests of Rational Expectations." *Journal of the European Economic Association*, 22, 2700–2747.

Cunha, F., I. Elo, and J. Culhane (2022): "Maternal subjective expectations about the technology of skill formation predict investments in children one year later," *Journal of Econometrics*, 231, 3–32, annals Issue: Subjective Expectations & Probabilities in Economics.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010): "Estimating the technology of cognitive and noncognitive skill formation," *Econometrica*, 78, 883–931.

Dasgupta, D. and A. Saha (2022): "Perceptions, Biases, and Inequality," *Journal of Economic Behavior and Organization*, 202, 198–210.

DellaVigna, S., J. A. List, and U. Malmendier (2012): "Testing for Altruism and Social Pressure in Charitable Giving," *Quarterly Journal of Economics*, 127, 1–56.

DeVellis, R. F. (2016): *Scale Development: Theory and Applications*, Sage Publications, 4 ed.

D'Haultfoeuille, X., C. Gaillac, and A. Maurel (2021): "Rationalizing Rational Expectations: Characterizations and Tests." *Quantitative Economics*, 12, 817–842.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," *Journal of the European Economic Association*, 9, 522–550.

Dominitz, J. and C. F. Manski (1996): "Eliciting Student Expectations of the Returns to Schooling," *Journal of Human Resources*, 31, 1–26.

——— (1997): "Using Expectations Data To Study Subjective Income Expectations," *Journal of the American Statistical Association*, 92, 855–867.

Duflo, E., R. Glennerster, and M. Kremer (2007): "Using Randomization in Development Economics Research: A Toolkit," in *Handbook of Development Economics*, ed. by T. P. Schultz and J. Strauss, Elsevier, vol. 4, 3895–3962.

Dutz, D., I. Huitfeldt, S. Lacouture, M. Mogstad, A. Torgovitsky, and W. van Dijk (2025): "Selection in Surveys: Using Randomized Incentives to Detect and Account for Nonresponse Bias." Working paper.

Dziewulski, P., J. Lanier, and J. K.-H. Quah (2024): "Revealed Preference and Revealed Preference Cycles: A Survey," *Journal of Mathematical Economics*, 113, 102023.

Evans, G. W. and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics*, Princeton, NJ: Princeton University Press.

Field, E., R. Pande, N. Rigol, S. Schaner, and C. Troyer Moore (2021): "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms," *American Economic Review*, 111, 2342–2376.

Fiori, G. and F. Scoccianti (2023): "The Economic Effects of Firm-Level Uncertainty: Evidence Using Subjective Expectations," *Journal of Monetary Economics*, 140, 92–105.

Fisman, R., S. Kariv, and D. Markovits (2007): "Individual Preferences for Giving," *American Economic Review*, 97, 1858–1876.

Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, 40, 351–401.

Freyberger, J. (2024): "Normalizations and Misspecification in Skill Formation Models," *arXiv preprint arXiv:2104.00473*.

Fudenberg, D. and D. K. Levine (1998): *The Theory of Learning in Games*, Cambridge, MA: MIT Press.

Giannola, M. (2023): "Parental Investments and Intra-household Inequality in Child Human Capital: Evidence from a Survey Experiment," *The Economic Journal*, 134, 671–727.

Gilboa, I., D. Samet, and D. Schmeidler (2004): "Utilitarian Aggregation of Beliefs and Tastes," *Journal of Political Economy*, 112, 932–938.

Giustinelli, P. (2023): "Chapter 7 - Expectations in education," in *Handbook of Economic Expectations*, ed. by R. Bachmann, G. Topa, and W. van der Klaauw, Academic Press, 193–224.

Giustinelli, P., C. F. Manski, and F. Molinari (2019): "Precise or Imprecise Probabilities? Evidence from Survey Response on Late-onset Dementia," Tech. rep., NBER.

——— (2022): "Tail and center rounding of probabilistic expectations in the Health and Retirement Study," *Journal of Econometrics*, 231, 265–281, annals Issue: Subjective Expectations and Probabilities in Economics.

Giustinelli, P. and M. D. Shapiro (2024): "SeaTE: Subjective Ex Ante Treatment Effect of Health on Retirement," *American Economic Journal: Applied Economics*, 16, 278–317.

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

——— (1972): "Structural equation methods in the social sciences," *Econometrica*, 40, 979–1001.

Goldberger, A. S. and O. D. Duncan, eds. (1973): *Structural Equation Models in the Social Sciences*, New York: Seminar Press.

Goldberger, A. S. and K. G. Jöreskog (1975): "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable," *Journal of the American Statistical Association*, 70, 631–639.

Guiso, L. and M. Paiella (2008): "Risk Aversion, Wealth, and Background Risk," *Journal of the European Economic Association*, 6, 1109–1150.

Guiso, L., P. Sapienza, and L. Zingales (2004): "The Role of Social Capital in Financial Development," *American Economic Review*, 94, 526–556.

——— (2008): "Trusting the Stock Market," *Journal of Finance*, 63, 2557–2600.

Harsanyi, J. C. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, 309–321.

Henzel, S. (2021): "Subjective Uncertainty, Expectations, and Firm Behavior," *ifo Working Paper Series*.

Hurd, M. D. (2009): "Subjective Probabilities in Household Surveys," *Annual Review of Economics*, 1, 543–562.

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.

Koşar, G. and C. O'Dea (2023): "Expectations Data in Structural Microeconomic Models," in *Handbook of Economic Expectations*, ed. by G. Topa, O. Coibion, Y. Gorodnichenko, and R. Kamdar, Elsevier, 647–675.

Kotlarski, I. I. (1967): "On Characterizing the Gamma and the Normal Distribution," *Pacific Journal of Mathematics*, 20, 69–76.

Lancaster, T. and A. Chesher (1983): "An Econometric Analysis of Reservation Wages," *Econometrica*, 51, 1661–76.

Lareau, A. (2011): *Unequal Childhoods: Class, Race, and Family Life*, Berkeley, CA: University of California Press, 2nd ed.

Lillard, L. A. and Y. Weiss (1979): "Components of Variation in Panel Earnings Data: American Scientists, 1960–70," *Econometrica*, 47, 437–454.

Lillard, L. A. and R. J. Willis (1978): "Dynamic Aspects of Earnings Mobility," *Econometrica*, 46, 985–1012.

Lovell, M. C. (1986): "Tests of the Rational Expectations Hypothesis." *American Economic Review*, 76, 110–124.

MaCurdy, T. E. (1982): "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," *Journal of Econometrics*, 18, 83–114.

Malmendier, U. and S. Nagel (2011): "Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?" *Quarterly Journal of Economics*, 126, 373–416.

Manski, C. F. (1990): "The Use of Intentions Data to Predict Behavior: A Best-Case Analysis," *Journal of the American Statistical Association*, 85, 934–940.

——— (2004): "Measuring expectations," *Econometrica*, 72, 1329–1376.

Meango, R. (2023): "Using Probabilistic Stated Preference Analyses to Understand Actual Choices," ArXiv preprint.

Messick, S. (1995): "Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning," *American Psychologist*, 50, 741–749.

Miller, R. A. (1984): "Job Matching and Occupational Choice," *Journal of Political Economy*, 92, 1086–1120.

Mohnen, P., B. H. Hall, J. Mairesse, and P. Mohnen (2008): "The Impact of Perceived Expectations and Uncertainty on Firm Investment," *Review of Economics and Statistics*, 90, 66–86.

Méango, R. and F. Poinas (2023): "The (Option-) Value of Overstaying." Tech. Rep. 10536, CESifo Working Paper.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85, 1313–1326.

Nickell, S. J. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417–1426.

Norris Keiller, A., A. de Paula, and J. Van Reenen (2024): "Production Function Estimation Using Subjective Expectations Data," Working Paper 32725, National Bureau of Economic Research.

Olley, S. G. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Pantano, J. and Y. Zheng (2013): "Using Subjective Expectations Data to Allow for Unobserved Heterogeneity in Hotz-Miller Estimation Strategies," SSRN Working Paper, available at SSRN: https://ssrn.com/abstract=2129303 or http://dx.doi.org/10.2139/ssrn.2129303.

Putnam, R. D. (2015): *Our Kids: The American Dream in Crisis*, New York: Simon and Schuster.

Schennach, S. M. (2016): "Recent advances in the measurement error literature." *Annual Review of Economics*, 8, 341–377.

Schulhofer-Wohl, S. (2011): "Heterogeneity and Tests of Risk Sharing," *Journal of Political Economy*, 119, 925–958.

Scur, D., R. Sadun, J. Van Reenen, R. Lemos, and N. Bloom (2021): "The World Management Survey at 18: Lessons and the Way Forward," *Oxford Review of Economic Policy*, 37, 231–258.

Stantcheva, S. (2022): "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible," Working Paper 30527, NBER.

Stinebrickner, R. and T. R. Stinebrickner (2014a): "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model," *Journal of Labor Economics*, 32, 601–644.

——— (2014b): "A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout," *Review of Economic Studies*, 81, 426–472.

Tobin, J. (1959): "On the predictive value of consumer intentions and attitudes," *The review of economics and statistics*, 1–11.

Todd, P. E. and K. I. Wolpin (2006): "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, 96, 1384–1417.

Townsend, R. M. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591.

van der Klaauw, W. (2012): "On the Use of Expectations Data in Estimating Structural Dynamic Choice Models," *Journal of Labor Economics*, 30, 521–554.

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, 21–42.

Varian, H. R. (1982): "The Nonparametric Approach to Demand Analysis," *Econometrica*, 50, 945–973.

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

——— (2018): "Preference for the workplace, investment in human capital, and gender." *Quarterly Journal of Economics*, 133, 457–507.

——— (2021): "Human Capital Investments and Expectations about Career and Family," *Journal of Political Economy*, 129, 1361–1424.

Wolpin, K. I. (1999): "Commentary on "Analysis of Choice Expectations in Incomplete Scenarios" by C.F. Manski," *Journal of Risk and Uncertainty*, 19, 67–69.

——— (2013): *The Limits of Inference without Theory*, Cambridge, MA: MIT Press.