The Dynamic Properties of Economic Preferences

Nicolás Salamanca*

April 2018

Abstract

The time-stability of preferences is a crucial and ubiquitous assumption in economics, yet to date there is no method to test its validity. Based on a model of the dynamics of individual preferences, I develop a simple method to test this assumption. Time-persistence in preferences is captured via an autoregressive parameter that accounts for observable characteristics and is unattenuated by measurement error, which forms the basis of the test. The method also estimates the variance of persistent shocks to latent preferences, which measures unobserved heterogeneity, and preference measurement error. I illustrate the use of this method by testing the stability of risk aversion and patience using micro-level data, and find that patience is time-stable but risk aversion is not. However, change very slowly over time. This method provides researchers with a simple tool to properly test the assumption on preference stability, and to measure the degree of preference changes due to observable and unobservable factors.

JEL classification: D01; D03; C18

Keywords: stability of preferences; risk aversion; patience; shock persistence; measurement error

---

*Melbourne Institute: Applied Economic & Social Research, The University of Melbourne; ARC Centre of Excellence for Children and Families over the Life Course; and IZA. Level 5, Business and Economics Building, 111 Barry Street, Victoria 3010, Australia. Email: n.salamanca@unimelb.edu.au Tel: (+61) 038 344 2863. I have received helpful comments from Deborah Cobb–Clark, Denis de Crombrugghe, Jan Feld, Alex de Gendre, Bart Golsteyn, Dan Hamermesh, Simon Georges-Kot, Stefi Schurer, Benedikt Vogt, Jongsay Yong, Ulf Zölitz, and participants in various seminars and conferences. An earlier version of this paper circulated as "A structured approach to testing the stability of individual preferences." This research was partly supported by the Australian Research Council Centre of Excellence for Children and Families over the Life Course (project number CE140100027). The Centre is administered by the Institute for Social Science Research at The University of Queensland, with nodes at The University of Western Australia, The University of Melbourne and The University of Sydney. In this paper use is made of data of the DNB Household Survey.

1 Introduction

Economists often assume that individual economic preference parameters, such as risk aversion or the rate of time discounting, remain constant over time. This assumption is crucial for most economic analyses. For example, the time-stability of economic preference primitives is often implicitly assumed when mapping cardinal utility into ordinal preferences. The stability of this mapping is central for the derivation of comparative statics and the ensuing welfare analyses, since unstable preferences imply a violation *ceteris paribus*. In finance, time-stable preferences are also a key identifying assumption, most notably in the Capital Asset Pricing Model (CAPM) and most of its variants. Time-varying risk preferences could generate unstable demands for risky assets, increasing the implied market risk aversion and biasing estimates of the market risk premia. In household economics, preference stability is often assumed for individual household members, and recent work shows that this is sufficient to identify budget shares in non-unitary household bargaining models (Valente and Sokullu, 2018). Stable preferences is also a crucial (and mostly implicit) assumption underlying the identification of parameters in structural models, and in the identification of policy–relevant treatment effects in reduced–form estimates.

Yet it is easy to imagine that preferences can change over time. People could change the way they feel about risk as they grow older, and as they gather different life experiences (Malmendier and Nagel, 2011; Sahm, 2012; Schurer, 2015; Dohmen et al., 2017; Hanaoka et al., forthcoming). These factors could also affect how people feel about the future (Krupka and Stephens, 2013). And a host of other factors, largely unobserved to researchers and potentially unobserved to people themselves, could impact their preferences in a persistent manner. In spite of the growing evidence on events that change preferences, there is no tool to assess the degree to which latent preferences gradually change over time, or to test their time-stability.

I fill this gap by developing a model for the dynamics of individual economic preferences and a simple way to estimate the model's key parameters, which provide a way to test the time-stability of preferences. In Section 2 I show how the model captures the evolution of latent preferences as a first–order autoregressive process, linking current preferences to past preferences via a single autoregressive parameter. This autoregressive parameter can be seen as an adjusted test–retest correlation—unattenuated by measurement error and net of predictable changes in preferences by observable characteristics—which quantifies the rate of change in latent preferences, and can be used to formally test their time-stability. The model also provides two additional important parameter estimates. First, it provides estimates of the variance of persistent shocks to latent preferences. These shocks are changes in latent preferences that cannot be predicted by observable characteristics or by the natural change in preferences measured by the autoregressive parameter. The variance of these shocks measures the degree of unobserved heterogeneity in latent preferences and is crucial component of preference stability. Second, the model provides estimates of the measurement error in stated preferences and corrects all estimates for its presence using a standard instrumental variable approach.

In Section 3 I illustrate the use of this model by testing the time-stability of risk aversion and patience using up to 15 years of data from the Dutch National Bank Household Survey (DHS). In these data, a formal test rejects the null hypothesis of time-stability for risk aversion but not for patience. However, the autoregressive parameter estimates indicate that both risk aversion and patience change very slowly over time, so slowly that they can be considered as stable over the span of a few years. I also document three other important findings. First, the variance of persistent shocks to both latent preferences are tiny, indicating that shocks unforecastable by researchers are unlikely to be important drivers of preference instability in these data. Second, I show that *changes* in preferences are not well predicted by observable individual characteristics but preferences *levels* are. This aligns well with previous findings in the literature (e.g., Sahm, 2012; Meier and Sprenger, 2015). Third, I estimate that over 30 percent of the variation in the measures of risk aversion and patience stems from measurement error, highlighting the importance of a method that can properly account for it in its estimates.

In Section 4 I present a number of extensions that *i*) discuss the stability of preferences over long horizons; *ii*) develop an alternative test for the time-stability of preferences that does not directly rely on the autoregressive parameter in the model; *iii*) explore ways of identifying parameter estimates under less restrictive data conditions; *iv*) describe the heterogeneity in the time-stability of preferences across several sub-groups; and *v*) simulate the potential bias introduced by using "stale" measures of preferences for predicting behavior (e.g., Einav et al., 2012, p. 2617). In Section 5 I conclude.

This paper makes an important methodological contribution, and a related empirical contribution. The methodological contribution is the development of a tool that can be used to support or disprove the often-made but rarely-warranted assumption of preference stability, which is central to most economic analyses.[1] If preferences are found to be time-changing, the method also provides researchers with a measure of the rate of change in preferences

[1] The most recent example of the growing interest in the stability of economic preferences is the special issue of the *Journal of Economic Psychology* (vol. 60, June 2017) on the topic. Amongst other papers in this special issue, Golsteyn and Schildberg-Horisch (2017) discusses the challenges for analysing the stability of preferences in economics and psychology, and highlights areas where further progress can be made; Hardardottir (2017) analyzes the long–term stability and determinants of time preferences, also using DHS data, and Duersch et al. (2017) investigates the intertemporal stability of ambiguity preferences. None of the papers issue, however, proposes a formal way to test the time-stability of the preferences they analyze.

attributable to gradual changes over time, changes due to observed factors, and changes due to unobserved shocks. This can help researchers assess how serious is the preference instability problem in their setting. If applied to various candidate preference measures, the method can also determine which of them is measured with less noise. But, perhaps more importantly for applied researchers, this method is easy to implement in terms of both data and technical requirements. The companion Stata code to this paper, `dynreg`, is available upon request.

The empirical contribution of this paper can be best viewed within the rapidly-growing literature that estimate the determinants of risk and time preferences. For risk preferences, Malmendier and Nagel (2011) show that past experienced stock returns affect people's risk aversion. Hoffmann et al. (2013) and Guiso et al. (2018) find similar relation triggered by the financial crisis of 2007. Sahm (2012) shows that age and macroeconomic conditions also correlate with risk aversion, but that persistent differences across individuals account for the majority of the variation. Dohmen et al. (2016) document relatively high test-retest correlations in risk preferences in Germany and Ukraine, yet also see that risk preference changes correlate with macroeconomic conditions. Bucciol and Miniaci (2018) show that risk aversion changes with both macroeconomic conditions and subjective risk exposure in people's past portfolio choices. Hanaoka et al. (forthcoming) document persistent changes in risk preferences after the 2011 Great East Japan Earthquake. For time preferences, Meier and Sprenger (2015) show that intertemporal discount rates generally do not change over a two-year period, and that the changes that do occur are unrelated to people's sociodemographic characteristics. Krupka and Stephens (2013), on the other hand, show that their time preference measures respond to the macroeconomic conditions faced by people, and therefore that time preference measures are more likely capturing the market interest rate faced by people than their time preferences.

The main empirical results of this paper are broadly consistent with those of the literature above, yet I provide the first formal test of the time-stability of risk and time preferences, and provide estimates of several additional parameters which further our understanding of the way these preferences evolve over time.

2 A Model of Time-Varying Preferences

2.1 Model

Abstracting from any specific preference, let person i's measured level of preference P at time t be P_{it}. Let P^*_{it} be the corresponding latent individual preference. I model the dynamics of preference P via the following two equations:

$$P_{it} = P_{it}^{*} + \varepsilon_{it} \tag{1}$$

$$P_{it}^{*} = \beta' X_{it} + \alpha P_{i,t-1}^{*} + \eta_{it} \tag{2}$$

Equation (1) defines the observed preference P_{it} as the sum of the latent preference P_{it}^{*} and a measurement error term ε_{it}. This measurement error is assumed to be classical — i.e., homoscedastic, independent of P_{it}^{*}, serially uncorrelated, and with well-defined second moments — although autocorrelated measurement error can be easily accommodated (see Section 2.2).

Equation (2) defines the evolution of the latent preference P_{it}^{*} as an autoregressive process of order one with a drift $\beta' X_{it}$.[2] For simplicity, I assume that the linear projection of X_{it} on P_{it}^{*} is a good approximation of the impact of individual characteristics on preferences, making β a sufficient statistic for describing the effect of observable characteristics on preference P. The drift, $\beta' X_{it}$, allows preferences to tend towards a conditional mean level determined by individual characteristics, effectively letting preferences "converge" towards different values for observably different people.[3] The error term η_{it} represents the idiosyncratic shocks to the latent preference P_{it}^{*}, assumed to be independent, homoscedastic, and with well-defined second moments.

Four parameters in this model provide useful information about the dynamics of preference P. The autoregressive parameter, α, identifies the time-stability of P_{it}^{*} — that is, how close is the current latent preference to the past latent preference. The parameter β identifies how individual characteristics determine the level that preferences naturally tend to. The variance of the idiosyncratic shocks to preferences, σ_{η}^{2}, identifies the severity of variations in latent preferences once individual characteristics have been accounted for. And the variance of the measurement error, σ_{ε}^{2} identifies the noise in the measurement of preference P.

To estimate the model's α and β parameters, begin by replacing (1) in (2) to obtain:

$$P_{it} = \beta' X_{it} + \alpha P_{i,t-1} + \{\eta_{it} + \varepsilon_{it} - \alpha \varepsilon_{i,t-1}\} = \beta' X_{it} + \alpha P_{i,t-1} + \upsilon_{it}, \tag{3}$$

and note that preferences in this equation are all observable, though the composite error term υ_{it} is correlated with $P_{i,t-1}$. Section 2.2 shows how solve this endogeneity issue and obtain

[2] More complicated autoregressive preference structures, such as an AR(2) process, can also be accommodated. However, The interpretation of the AR(2) parameters is not clear, and empirically they add nothing to the AR(1) model.

[3] One could simply write the model in terms of *de-drifted* preferences. Carroll and Samwick (1997) take this approach when calculating the variance of permanent income. If the drift is correlated with the lagged preferences, however, de-drifting the process without including lagged preferences as an additional explanatory variable will result in biased estimates of the drift parameters.

consistent estimates of α and β..

To derive expressions for the variance of η_{it} and ε_{it} it is simpler to work with the de-drifted preference, $\tilde{P}^*_{it} = P^*_{it} - \beta' X_{it}$, which can be easily constructed from the data given consistent estimates of β from Equation (3). Taking the k^{th} difference of $\tilde{P}_{it}$ and replacing recursively for $\tilde{P}^*_{i,t+k} = \alpha \tilde{P}^*_{i,t+k-1} + \eta_{i,t+k}$ using (2) yields:

$$
\begin{aligned}
\tilde{P}_{i,t+k} - \tilde{P}_{it} &= \tilde{P}^*_{i,t+k} + \varepsilon_{i,t+k} - \tilde{P}^*_{it} - \varepsilon_{it} \\
&= \alpha \tilde{P}^*_{i,t+k-1} + \alpha\beta' X_{i,t+k-1} + \eta_{i,t+k} + \varepsilon_{i,t+k} - \tilde{P}^*_{it} - \varepsilon_{it} \\
&= \alpha\{\alpha \tilde{P}^*_{i,t+k-2} + \alpha\beta' X_{i,t+k-2} + \eta_{i,t+k-1}\} + \alpha\beta' X_{i,t+k-1} + \eta_{i,t+k} + \varepsilon_{i,t+k} - \tilde{P}^*_{it} - \varepsilon_{it} \\
&\vdots \\
&= (\alpha^k - 1)\tilde{P}^*_{it} + \sum_{j=1}^{k} \alpha^j \beta' X_{i,t+k-j} + \sum_{j=0}^{k-1} \alpha^j \eta_{i,t+k-j} + \varepsilon_{i,t+k} - \varepsilon_{it}.
\end{aligned}
$$

Replacing for $\tilde{P}^*_{it} = \tilde{P}_{it} - \varepsilon_{it}$ and rearranging terms results in:

$$
\tilde{P}_{i,t+k} - (\alpha^k)\tilde{P}_{it} = \sum_{j=1}^{k} \alpha^j \beta' X_{i,t+k-j} + \sum_{j=0}^{k-1} \alpha^j \eta_{i,t+k-j} - \alpha^k \varepsilon_{it} + \varepsilon_{i,t+k}. \tag{4}
$$

The conditional variance of Equation (4) provides an expression that can be used to estimate the remaining parameters in the model. Note that the left-hand side of (4) is expressed only in terms of $\tilde{P}_{i,t}$, which is observable, and α, for which there is consistent estimator. We can therefore obtain estimates of its variance, $Var[\tilde{P}_{i,t+k} - (\alpha^k)\tilde{P}_{it}]$ for every k as long as there is data for periods t and $t+k$. The first term on the right-hand side of (4) shows that the variation due to individual characteristics can be accounted for by conditioning on $\bar{X}_{i,t+k} = (k-1)^{-1}\sum_{j=1}^{k} X_{i,t+k-j}$, the averages of individual characteristics up to time $t+k-1$. Taking the variance of (4) conditional on $\bar{X}_{i,t+k}$ yields:

$$
Var[\tilde{P}_{i,t+k} - (\alpha^k)\tilde{P}_{it} | \bar{X}_{i,t+k}] = \sigma^2_\eta \sum_{j=1}^{k} \alpha^{2j} + \sigma^2_\varepsilon(\alpha^{2k} + 1). \tag{5}
$$

Through Equation (5) we can identify the variances of shocks to latent preferences and preference measurement error. The left-hand side of (5) is all observed. The right-hand side if (5) includes the terms $\sum_{j=1}^{k} \alpha^{2j}$ and $(\alpha^{2k} + 1)$, for which we have consistent estimators, and the variances σ^2_η and σ^2_ε, which are the parameters of interest. These parameters can be obtained as the solutions to a two-unknown linear equation system, an extension of the technique used to identify the variance of permanent income by Carroll and Samwick (1997). For example, replacing α by its consistent estimator, $\hat{\alpha}$, and taking data for two lags, $k=1$,2 we can solve the system

$$Var[\tilde{P}_{i,t+1}-\hat{\alpha}\tilde{P}_{it}|\bar{X}_{i,t+1}]=\sigma_\eta^2\hat{\alpha}^2+\sigma_\varepsilon^2(\hat{\alpha}^2+1)$$
$$Var[\tilde{P}_{i,t+2}-(\hat{\alpha}^2)\tilde{P}_{it}|\bar{X}_{i,t+2}]=\sigma_\eta^2(\hat{\alpha}^2+\hat{\alpha}^4)+\sigma_\varepsilon^2(\hat{\alpha}^4+1)$$

for σ_η^2 and σ_ε^2. This yields estimators for $\hat{\sigma}_\eta^2$ and $\hat{\sigma}_\varepsilon^2$ which, together with $\hat{\alpha}$ and $\hat{\beta}$ complete the model's parameter estimates.

2.2 Estimation

Estimates of α and β can be directly obtained from (3). However, an ordinary least squares will yield inconsistent estimates since $P_{i,t-1}$ and $\varepsilon_{i,t-1}$ are correlated. To obtain consistent estimates, I exploit the fact that P_{ij} is uncorrelated with the composite error term in (3) for $j<t-2$, or equivalently that the moment conditions $E[P_{ij}\cdot\upsilon_{it}]=0\forall j=0,...,t-2$ identify the parameters of interest. These are straightforward moment choices given the time structure of the problem, and the same moment conditions implied in an instrumental variable (IV) approach, where $P_{i,t-1}$ is instrumented by further lags of P. Therefore, equation (3) can be easily estimated via IV.

There are several other moment conditions that can be used for identification, some of which are explored in Section 4.3. Different moment conditions can be used to balance data requirements and the strictness of the underlying assumptions for the validity of the instruments. For example, serial correlation in ε_{it} does not make the method unfeasible; it merely restricts the lag structure of the valid instruments available for estimation. If $\varepsilon_{i,t-1}$ follows an AR(3) process, the IV estimation will have to *i*) exploit lags of P_{ij} starting at $j=t-3$, or *ii*) exploit other moment conditions that hold given this autocorrelation structure in the error term. Other data restrictions and model extensions will imply different valid instruments, yet many configurations will be admitted in the general IV estimation framework.

The variance of the idiosyncratic shocks (σ_η^2) and the measurement error (σ_ε^2) can be estimated solving the 2-equations 2-unknown system in (5) for two arbitrary periods. However, a more efficient estimator combines the information of all valid k–lengths via a non-linear regression, obtaining the variances as non-linear least squares parameters constrained to be positive. Finally, since P_{it} can easily be standardized to have unit variance, the noise-to-signal ratio, a comparable metric of the amount of measurement error in preferences across models, can be estimated as $s=\sigma_\varepsilon^2/(1-\sigma_\varepsilon^2)$ (Cameron and Trivedi, 2005, p. 903).

2.3 Interpretation, and Inference

The main goal is to provide a formal test of the time-stability of preference P. In this model such test is for $\hat{\alpha}$ to be equal to one; that is, whether the autoregressive parameter in preferences indicates that preferences remain as they were in the last period. Interpretation of $\alpha = 1$, however, is not straightforward. If viewed as a test-retest correlation, $\alpha = 1$ would imply a perfect correlation over time and thus complete stability. However, from a time series perspective, $\alpha = 1$ implies that any idiosyncratic shock to preferences will be infinitely-lived (i.e., a unit root) and thus preferences would be, in a way, unstable. Stability thus needs to be judged by jointly looking at α and σ_η^2, the autoregressive parameter and the idiosyncratic shock variance. Hence the importance of having a method that jointly estimates both. Given the evidence on the size of σ_η^2 in Section 3.2, it seems natural to view $\alpha = 1$ as a sign of preference stability, which aligns with a classical microeconometrician's interpretation. This needs not hold for larger values of σ_η^2, though.[4]

Statistical inference for α is straightforward, but some remarks are worth making. As α approaches unity, testing for $\hat{\alpha} = 1$ becomes increasingly similar to a panel data unit root test. The similarity suggests that, should preferences have a unit root, $\hat{\alpha}$ could have a distribution that degenerates too fast as the number of individuals, time periods, or observations tends to infinity. This phenomenon, called super-consistency, makes t-statistics unsuitable for inference. Worse, the estimator could have a distribution whose asymptotic behavior depends on the rate at which the number of individuals and the time periods tend to infinity.

However, when testing the stability of preferences in this setting, information accumulates by increasing the number of individuals, not each individual's lifespan. This means that the relevant asymptotics for inference in this model are large N and fixed T. Harris and Tzavalis (1999) show that with large N and fixed T asymptotics, the limiting distribution of $\sqrt{N}(\hat{\alpha}_{OLS} - 1) \xrightarrow{d} N(0, 2\{T(T-1)\}^{-1})$. I use their test for inference on $\hat{\alpha}$ noting that their derivation is only valid for balanced panels so I need to accommodate my tests to this restriction (see Table 3). A minor issue is that the efficiency loss in IV compared to OLS implies that standard errors can be somewhat larger for my estimates, making the panel unit root tests conservative. More importantly, however, the estimator $\hat{\alpha}$ is (super-)consistent regardless on which way the standard errors are calculated and used. This feature is useful for constructing the additional test for the time–stability of preferences described in Section 4.2.

A final consideration is the potential existence of confounding unobserved individual heterogeneity. In this setting, confounding heterogeneity would take the form of individual

[4]Note that $\alpha = 1$ implies time-stability but also equilibrium-*in*stability, since shocks to preferences are infinitely-lived. These two concepts of stability are, thus, irreconcilable. I chose to interpret the models estimates under the former view, since it is the one traditionally taken by researchers.

unobserved differences in preferences that, if correlated with lagged preferences or individual characteristics, could bias the estimates. The standard way to account for this heterogeneity is to eliminate its time-invariant individual-specific component via fixed or random effects (e.g., Sahm, 2012). In this model, however, this heterogeneity is the identifying variation of the estimates, and eliminating it makes no sense. To see this more clearly, consider the scenario where preferences are time-stable, and thus $P^*_{it} = P^*_i$. In this model, latent preferences would be fully captured by time-invariant individual heterogeneity, and in any panel model that eliminates it, such as Arellano and Bond (1991), the autoregressive parameter α will be unidentified under the null hypothesis. Therefore, I do not consider these models as feasible estimators for the purposes of this paper. Mixed effect models, on the other hand, can accommodate time-invariant individual unobserved heterogeneity uncorrelated to lagged preferences but partially correlated to individual characteristics (see Mundlak (1978) for their development and Sahm (2012) for an application). Since the coefficients of controls are not the focus of this paper, and such models would not aid in the identification of the autoregressive parameter, I do not consider them in this paper. Instead, I focus on estimating $\hat{\alpha}$, $\hat{\sigma}^2_\eta$ and $\hat{\sigma}^2_\varepsilon$ in the empirical exercise in the sext section.

3 The Dynamic Properties of Risk and Time Preferences

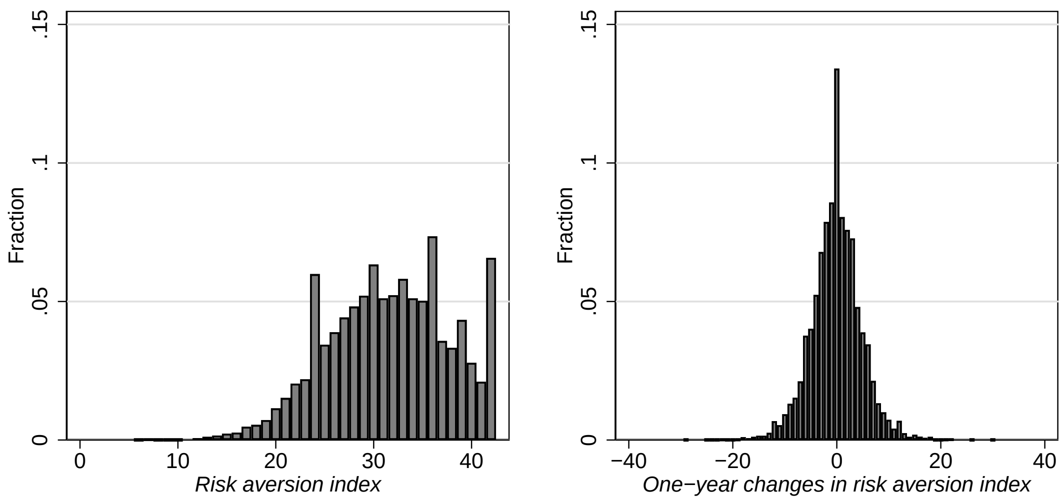
3.1 Data

In this section I use the model described above to estimate the dynamic properties and test the stability of risk and time preferences using data from the Dutch National Bank Household Survey (DHS).[5] The DHS is a yearly panel survey running on a representative sample of Dutch households since 1993. In the survey, every household member over the age of 16 answers a number of questions about the household organization, education, work situation and labor income, assets and liabilities, health, and economic and psychological concepts. On average, about 2,000 households (4,500 people) are interviewed each year, and each household stays in the panel for about seven and a half years on average.

I estimate the model's parameters for risk aversion and patience. The DHS measures these preferences by asking the respondents to rate their agreement with several statements for each preference on discrete ordinal scales. Risk aversion is measured through six items asked between 1994 and 2015. The items are mainly about financial risk attitudes and behavior so the index should be interpreted as measuring financial risk aversion. Patience is measured

[5]For a detailed description of the DHS, see Kapteyn and Teppa (2011) and Salamanca et al. (2016), or visit the CentERdata website at `www.centerdata.nl`.

Figure 1: Distribution of risk aversion levels and one-year changes in the DHS

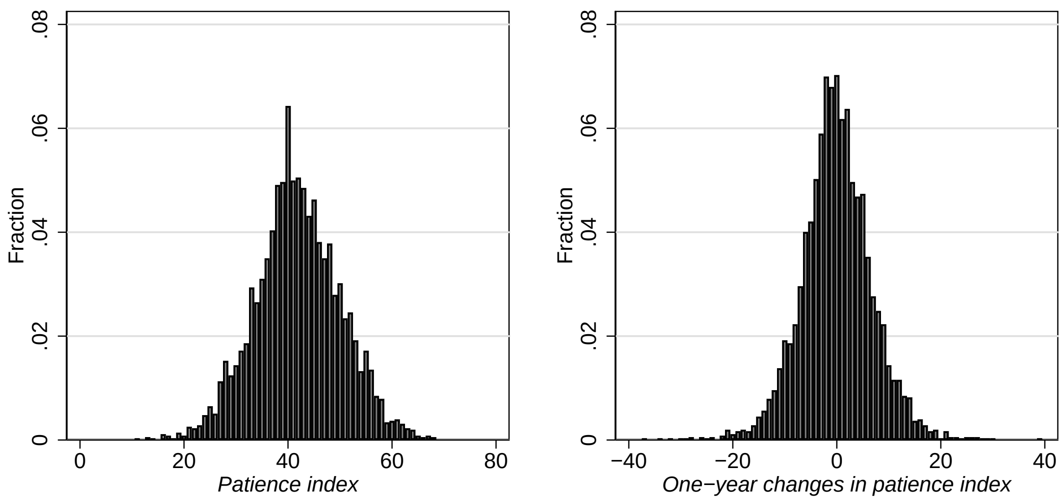


through 10 items asked between 1996 and 2007 which capture trade-offs between present and future rewards and pains, and general future-oriented thinking. These measures of risk aversion and patience have been behaviorally validated in several studies such as Warneryd (1996), Borghans and Golsteyn (2006), Kapteyn and Teppa (2011), and Salamanca et al. (2016).

I construct indices for both risk aversion and patience by reversing the items that measure each trait negatively and then adding up the trait's individual item scores. Figures 1 and 2 show the distribution of the constructed indices and of their one-year changes. There is substantial heterogeneity in both indices and in their changes, suggesting that the underlying preferences are potentially time-varying. However, part of the one-year change variability is surely due to measurement error. For all subsequent analyses I standardize both risk aversion and patience indices so they have zero mean and a standard deviation of one over their respective estimation samples. I use these indices as measures of P_{it}.

Table 1 provides details and summary statistics on each of the individual items and the indices measuring risk aversion and patience. My main estimation samples include 10,826 observations (2,964 individuals) for risk aversion, and 3,710 observations (1,473 individuals) for

Figure 2: Distribution of patience levels and one-year changes in the DHS



patience. I use only respondents between the ages of 16 and 80, who are either the household head, their partner, or a child, from households with positive yearly income but below €150,000, and with positive net wealth but below €1,000,000. These restrictions are imposed to improve the quality of the data used. Since the main focus of this paper is on the estimation of the autoregressive process and idiosyncratic shock variance in preferences, I include only a parsimonious set of control variables in X_{it}: household income, financial wealth, and number of children present, and individual occupation, age, gender, and marital status. Table A.1 in the Appendix shows summary statistics for these variables.[6]

[6]In any long-running panel survey such as this, there will be panel attrition and item non-response. In unreported analyses I verify that sample attrition is not related to changes in risk and time preferences (or to their leads and lags), which ensures that the estimates of α will not be biased by selective attrition. For item non-response, I include missing control categories for occupation and marital status, and select samples with non-missing values for the rest of the variables in the analysis.

Table 1: Summary statistics for the risk aversion and patience items

Item	Mean	S.D.	Percentile				
			Min	25^{th}	50^{th}	75^{th}	Max
Risk aversion ($N=10{,}826$)							
(a) I think it is more important to have safe investments and guaranteed returns, than to take a risk to have a chance to get the highest possible returns	5.14	1.75	1	4	6	7	7
(b) I would never consider investments in shares because I find this too risky	4.55	2.06	1	3	5	7	7
(c) If I think an investment will be profitable, I am prepared to borrow money to make this investment (r)	5.40	1.43	1	5	6	6	7
(d) I want to be certain that my investments are safe	5.87	1.51	1	5	7	7	7
(e) I get more and more convinced that I should take greater financial risks to improve my financial position (r)	5.15	1.66	1	4	5	7	7
(f) I am prepared to take the risk to lose money, when there is also a chance to gain money (r)	5.45	1.53	1	4	6	7	7
Sum of scores of risk aversion	31.56	6.25	6	27	32	36	42
Patience ($N=3{,}710$)							
(a) I think about how things can change in the future, and try to influence those things in my everyday life.	4.11	1.45	1	3	4	5	7
(b) I often work on things that will only pay off in a couple of years.	3.60	1.50	1	2	4	5	7
(c) I am only concerned about the present, because I trust that things will work themselves out in the future (r)	3.46	1.42	1	2	4	4	7
(d) With everything I do, I am only concerned about the immediate consequences (say a period of a couple of days or weeks) (r)	4.95	1.34	1	4	5	6	7
(e) I am ready to sacrifice my well-being in the present to achieve certain results in the future	4.18	1.34	1	3	4	5	7
(f) I think it is important to take warnings about negative consequences of my acts seriously, even if these negative consequences would only occur in the distant future	4.25	1.51	1	3	4	5	7
(g) I think it is more important to work on things that have important consequences in the future, than to work on things that have immediate but less important consequences	4.28	1.54	1	3	4	6	7
(h) In general, I ignore warnings about future problems because I think these problems will be solved before they get critical (r)	4.70	1.36	1	4	5	6	7
(i) I think there is no need to sacrifice things now for problems that lie in the future, because it will always be possible to solve these future problems later (r)	4.13	1.39	1	3	4	5	7
(j) I only respond to urgent problems, trusting that problems that come up later can be solved in a later stage (r)	4.22	1.43	1	3	4	5	7
Sum of scores of patience	41.88	8.23	11	37	42	47	68

Summary statistics for six risk aversion statements asked in 1996-2015, and 10 patience statements asked in 1998-2007. (r) indicates that the item was reversed for the sum of scores. Risk aversion statements have the following common heading: "The following statements concern saving and taking risks. Please indicate for each statement to what extent you agree or disagree. Please indicate on a scale from 1 to 7 to what extent you agree with the following statements, where 1 indicates 'totally disagree' and 7 indicates 'totally agree'." Patience statements have the following common heading: "To what extent do you agree or disagree with the following statements. If you really don't know, type 0 (zero).' Right below they show a numbered scale from 1 (labeled 'totally disagree') to 7 (labeled 'totally agree').

3.2 Main Results

Tables 2 shows the estimates of the autoregressive parameter, α, and variance estimates of the idiosyncratic shocks, σ_η^2, the measurement error, σ_ε^2, and the noise-to-signal ratio, s, for risk aversion and patience. The first two columns report biased OLS estimates of the autoregressive parameters (with and without adding controls to the model), for comparison. Columns (3) and (4) show the IV estimation that uses the second lag of preferences as instrument for the lagged preferences, and calculates the variance terms using the non-linear least squares described in Section 2.2.

Table 2: The dynamic properties of risk aversion and patience in the DHS

	OLS		IV	
	(1)	(2)	(3)	(4)
Dependent variable: Risk aversion (N=10,826)				
Lagged risk aversion (α)	0.699	0.670	0.970	0.966
	(0.010)	(0.010)	(0.009)	(0.010)
Idiosyncratic variance (σ_η^2)			0.048	0.048
Measurement error variance (σ_ε^2)			0.296	0.295
Noise-to-signal ratio (s)			0.420	0.419
Controls:	No	Yes	No	Yes
H_0: Controls jointly significant:	-	[0.001]	-	[0.605]
Dependent variable: Patience (N=3,710)				
Lagged patience (α)	0.647	0.636	0.984	0.985
	(0.017)	(0.017)	(0.019)	(0.019)
Idiosyncratic variance (σ_η^2)			0.000	0.000
Measurement error variance (σ_ε^2)			0.358	0.358
Noise-to-signal ratio (s)			0.557	0.557
Controls:	No	Yes	No	Yes
H_0: Controls jointly significant	-	[0.002]	-	[0.278]

Main parameters describing the dynamic properties of risk aversion and patience, estimated via OLS and IV. Second lag of preferences used as an instrument for the first lag in the IV regressions. Controls include income, financial wealth, age, occupation, gender, number of children, and marital status. Robust standard errors clustered at the individual level in parentheses. P-values in square brackets. σ_η^2, σ_ε^2, and s calculated via non-linear least squares using information from 2 lags and netting out X_{it}, as described in Section 2.2.

Columns (1) of Table 2 shows OLS autoregressive parameter estimates of 0.699 for risk aversion and 0.670 for patience. Both parameters are precisely estimated and are close to the highest one- and two-year test-retest correlations found for economic preferences (e.g., Dohmen et al., 2016), and also for personality traits at the life stage where personality is consider to be stable by psychologists (Roberts and DelVecchio, 2000). Correlations of this magnitude are generally considered to be high in the literature, yet the OLS coefficients are clearly far from one. Column (2) shows that accounting for potential correlates of risk aversion and patience barely changes this correlation.

Columns (3) and (4) report the same parameters estimated using IV, where the autoregressive parameter α is estimated by instrumenting lagged risk aversion and patience with their corresponding second lags. This correction consistently estimates all the parameters in the model, as described in Section 2.2. The IV estimates, at 0.966 for risk aversion and 0.985 for patience, are much closer to one than those in Columns (1) and (2) and are still precisely estimated.

Table 2 also shows estimates of the variance of idiosyncratic shocks to risk aversion and patience (σ_η^2), estimates of the variance of the measurement error of both preferences (σ_ε^2), and estimates of the noise-to-signal ratio (s) in each construct.[7] Recall that the variance of the risk preference and patience indices are normalized to one, so the estimated idiosyncratic shock variance of 0.048 for risk aversion is only around 5% of the overall variance in the preference index (which includes variance due to measurement error). This variance is remarkably low. For patience, the estimated variance is even smaller, only showing up in the fourth decimal digit, indicating that idiosyncratic shocks to patience are negligible.

The measurement error variance for risk aversion and patience, however, are 0.295 and 0.358, which mean that over 29 percent of the variation in risk preferences and over 35 percent of the variation in patience is due to measurement error. These high measurement error estimates naturally lead to high noise-to-signal ratios (0.419 and 0.557), which are consistent with the differences between the coefficients estimated via OLS and IV and indicate that a large part of the apparent year-to-year changes in risk aversion and patience are not due to changes in latent preferences.

The last rows in each panel of Table 2 shows that the joint significance test for the coefficients on the individual characteristics rejects the null of no significance for OLS estimates, but cannot reject the null for IV estimates. To see why this happens, note that the estimates of β are the partial effects of individual characteristics on preferences *conditional* on the natural year–to–year persistence of preferences as captured by the models' estimates of α. If α is estimated to be zero, β will reflect the marginal effects of X_{it} on the *level* of preferences. If α is estimated to be

[7]Cobb-Clark and Schurer (2013) show that the noise-to-signal ratio can also be estimated via structural equations systems. However, those methods usually impose strong distributional assumptions on the underlying preference and its covariance with other variables. Some of the distributional assumptions imposed, like normality, may not hold for some preferences (see, for example, the right tail in Figure 1). The method used in this paper makes no assumptions about the distribution of preferences, and can be adapted to account for more complex structures in measurement error.

one, β will reflect the marginal effects on the *changes* of preferences. (To see this, replace $\alpha = 1$ in (3) and subtract $P_{i,t-1}$ on both sides.) Since the estimates of α for risk aversion and patience are very close to one in the IV models, the non-rejection of F-Tests at the bottom of Table 2 indicate that individual characteristics are bad predictors of *changes* in preferences. And since the OLS estimates of α are much closer to zero, the rejection of F-Tests in these models indicate that individual characteristics are better predictors of *levels* in preferences. The fact that levels, but not changes, in preferences are well-predicted by observable characteristics is consistent with most of the previous findings in the literature, including Dohmen et al. (2011), Sahm (2012), Krupka and Stephens (2013), Dohmen et al. (2016), and Bucciol and Miniaci (2018).[8]

A test of the stability of risk and time preferences ($\alpha = 1$) using the estimates and standard errors reported on Table 2 would reject the null of stability for risk aversion, and would not be able to reject the null for patience. However, as discussed in Section 2.3, if estimates of α are super-consistent, it could be misleading to test for the time-invariance of preferences using these standard errors. A more appropriate method would use the balanced panel unit root test and asymptotic distributions developed in Harris and Tzavalis (1999). Table 3 shows these panel unit root tests for all possible balanced panels in the DHS sample. These tests have the null hypothesis that preferences are time-invariant ($\hat{\alpha} = 1$). To maximize the number of observations in each balanced panel, I *i*) regress each preference index on the individual characteristics and a complete set of year dummies, *ii*) follow Levin et al. (2002) in interpreting the residuals of that regression as a measure of preferences net of the cross-sectional influence of individual characteristics and common cross-sectional time effects, *iii*) construct balanced panels of size T which included all individuals with T adjacent residuals, and *iv*) calculate OLS and IV estimates of α, and their corresponding small-sample z-statistic from Harris and Tzavalis (1999), defined as $z = (\hat{\alpha} - 1)\sqrt{\frac{N}{2\{(T-1)(T-2)\}}}$.

The first two columns of Table 3 report the number of periods (T) and observations (N) in each of these balanced panels. The next three columns report the OLS estimates of the autoregressive parameter, the z-statistic, and the p-value of the Harris-Tzavalis panel unit root test. The last six columns report the IV estimates of the autoregressive parameter, the idiosyncratic and measurement error variances, the noise-to-signal ratio, the z-statistic, and the

[8]The unreported coefficients in this table reveal some familiar patterns. For OLS estimates, they describe a decrease in the level of risk aversion as wealth and income increase, a steeply–rising age pattern that predicts a peak around age 75, and females being substantially more risk-averse than males. The IV estimates suggest that only gender is a significant predictor of changes in risk aversion, capturing the fact that women are more likely to change their risk preferences between any two years. OLS and IV coefficients tell a similar story for patience; levels of patience are well predicted by observable characteristics but changes in patience are only predicted by retirement status and the number of children in the household.

Table 3:
Adaptation of the Harris and Tzavalis (1999) panel unit root tests on risk aversion and patience

		OLS			IV					
Panel size (T)	Obs.	α	z	p-value	α	σ_η^2	σ_ε^2	s	z	p-value
Risk aversion										
3	576	0.675	−5.511	0.001	0.903	0.187	0.206	0.260	−1.642	0.050
4	789	0.617	−10.766	0.001	0.929	0.135	0.242	0.320	−2.830	0.002
5	644	0.618	−11.813	0.001	0.914	0.162	0.263	0.357	−4.646	0.001
6	675	0.692	−11.204	0.001	0.896	0.139	0.234	0.306	−7.626	0.001
7	702	0.643	−14.906	0.001	1.009	0.000	0.316	0.463	0.865	0.194
8	602	0.629	−15.626	0.001	1.012	0.035	0.285	0.400	1.285	0.099
9	528	0.665	−14.234	0.001	0.983	0.008	0.314	0.458	−1.949	0.026
10	729	0.636	−19.634	0.001	0.942	0.051	0.273	0.375	−8.922	0.001
11	590	0.702	−15.062	0.001	1.014	0.010	0.250	0.333	2.110	0.017
12	341	0.677	−13.126	0.001	0.953	0.069	0.239	0.313	−6.194	0.001
13	516	0.699	−15.872	0.001	0.974	0.040	0.233	0.305	−4.641	0.001
14	195	0.835	−5.433	0.001	0.968	0.115	0.150	0.176	−3.819	0.001
15	182	0.797	−6.833	0.001	1.072	0.000	0.218	0.279	8.917	0.001
Patience										
3	414	0.588	−5.921	0.001	0.788	0.370	0.163	0.195	−3.045	0.001
4	486	0.662	−7.418	0.001	1.025	-	0.376	0.603	0.791	0.215
5	512	0.651	−9.638	0.001	0.957	0.058	0.342	0.520	−2.085	0.019
6	160	0.678	−5.748	0.001	0.926	0.065	0.259	0.349	−2.661	0.004
7	150	0.528	−9.066	0.001	1.072	-	0.362	0.567	3.100	0.001
8	21	0.469	−3.990	0.001	1.171	0.000	0.318	0.467	3.315	0.001
9	16	0.850	−0.026	0.490	-	-	-	-	-	-
10	45	0.417	−7.565	0.001	-	-	-	-	-	-

Harris and Tzavalis (1999) panel unit root tests for various balanced panels of risk aversion and patience. Cross-sectional means the effect of controls (income, financial wealth, age, occupation, gender, number of children, and marital status) subtracted from each preference to decrease their effect on the test, following Levin et al. (2002). z-statistics are calculated assuming an asymptotic variance of $2[(T-1)(T-2)]^{-1}$. σ_η^2, σ_ε^2, *and s calculated via non-linear least squares using information from 2 lags and netting out* X_{it}, *as described in Section 2.2.*

p-value of the Harris-Tzavalis test. The table shows, with few exceptions, remarkably consistent estimates of the OLS and IV autoregressive parameter for all panels with those reported in Table 2. Unsurprisingly, the Harris-Tzavalis test always rejects the null of time-stable preferences when based on OLS estimates. More interestingly, the test also rejects the null of stable preferences for risk aversion and for patience when based on IV estimates. This suggests that super-consistency could have led us to a type II error when testing the stability of patience based on the initial standard errors. However, I take these results with a grain of sand, since the estimates of the idiosyncratic variance are much less stable across balanced panels for both risk aversion and patience, and they are often larger than those in Table 2. This points to large gains of estimating this model's parameters in larger samples. Measurement error and noise-to-signal ratio estimates are far less erratic, though, and support the previous conclusion that both preferences are measured with substantial error, and patience more so than risk aversion.

Overall, the Harris-Tzavalis test indicate that we should reject the hypothesis that risk preferences and patience remain time-stable. However, the IV estimates of the time-variation of preferences suggest that even if preferences are not fully time-stable, they move very slowly from year to year. The degree to which this slow movement can effectively be characterized as time-instability is one of the issues discussed in the next section.

4 Extensions

In this section I present a number of extensions to the work above by *i*) discussing the stability of preferences over long horizons; *ii*) developing an alternative test for the time-stability of preferences that does not directly rely on the autoregressive parameter in the model; *iii*) exploring ways of identifying parameter estimates under less restrictive data conditions; *iv*) describing the heterogeneity in the time-stability of preferences across several sub-groups; and *v*) simulating the potential bias introduced by using "stale" measures of preferences for predicting behavior.

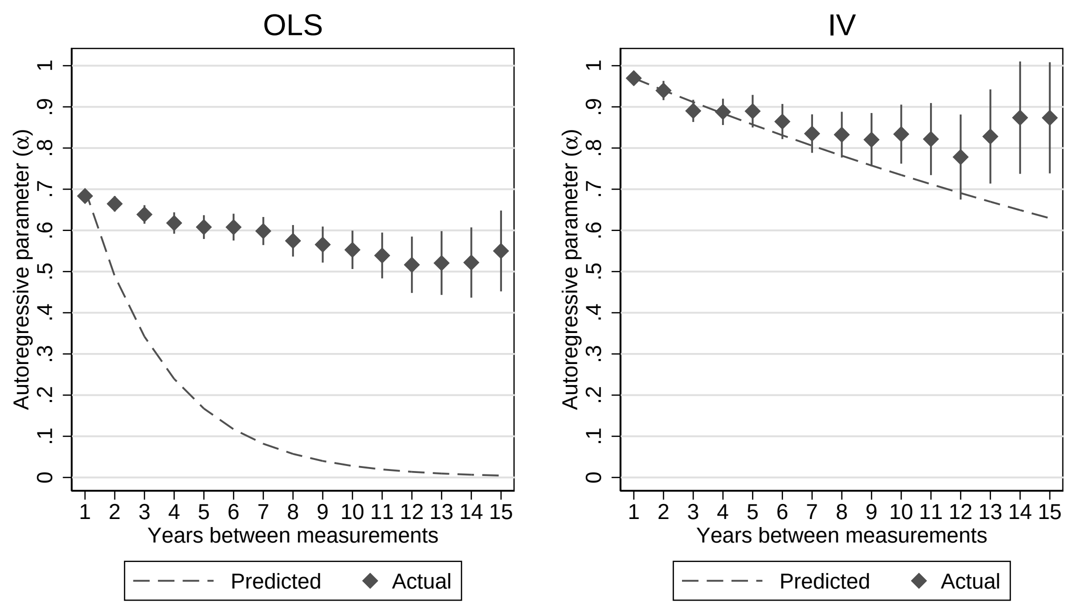
4.1 Dynamic Properties over different horizons

The estimates in Section 3 for both risk aversion and patience are all based on one–year horizon changes. However, estimating the models' parameters over longer periods of time is important for at least two reasons: it provides relevant information about the long-term stability of preferences, and it provides yet more evidence that OLS estimates are not informative of how stable preferences are.

To analyze long–term stability, we can re-estimate the model's parameters by relating current preferences to preferences measured two, three, four, and so on years in the past. I do this in the naïve OLS and the IV setting, always using the immediately previous lag as an instrument (e.g., for the 5–year horizon I regress preferences at time t on preferences at time $t-5$ via OLS, and do the same but instrument $t-5$ preferences with $t-6$ preferences for the IV estimate). These results can be compared to the implied autocorrelations predicted by our one–year estimates reported in Table 2. For example, the OLS estimates in Table 2 would predict the five-year autocorrelation in risk aversion to be $0.699^5 = 0.166$ whereas the IV estimates would predict it to be $0.973^5 = 0.841$.

Figures 3 through 4 illustrate these results. Performing this exercise for up to 15 years of differences in risk aversion and up to 10 years for patience shows that the OLS stability

Figure 3: Predicted vs actual autoregressive coefficients for risk aversion



predictions could easily lead to the wrong conclusion that preferences are time-varying even over relatively short periods of time. However, the IV estimates show that both risk aversion and patience remain relatively unchanging for over a decade.

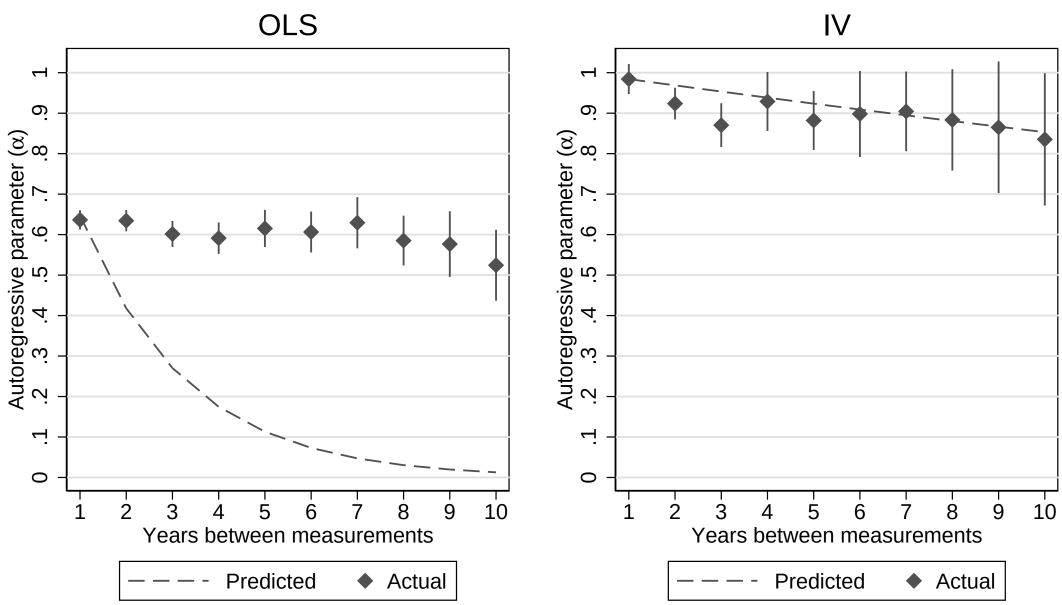
4.2 Variance-based evidence of stability

Due to the super-consistency issues discussed in Section 2.3, it would be useful to have a test for the time-stability of preferences that does not rely on the parameter estimates $\hat{\alpha}$. Here I develop such test by first noting that, if preferences are truly stable but measured with error, then $P_{it} = P_i^* + \varepsilon_{it}$.[9] Under the same assumptions made in Section 2, the variance of $P_{it} - P_{i,t-k}$ should remain constant for all k. If preferences change over time, however, this needs not be the case. This intuition can be formally built into a test for the time-invariance of preferences that does not rely on the standard t-statistic of α, and therefore works even under super-consistency.

Assuming that preferences evolve according to equations (1) and (2), and through a recursive

[9]When other researchers write about stable preferences, this is the model they usually have in mind. This model is explicitly written in Cobb-Clark and Schurer (2013).

Figure 4: Predicted vs actual autoregressive coefficients for patience



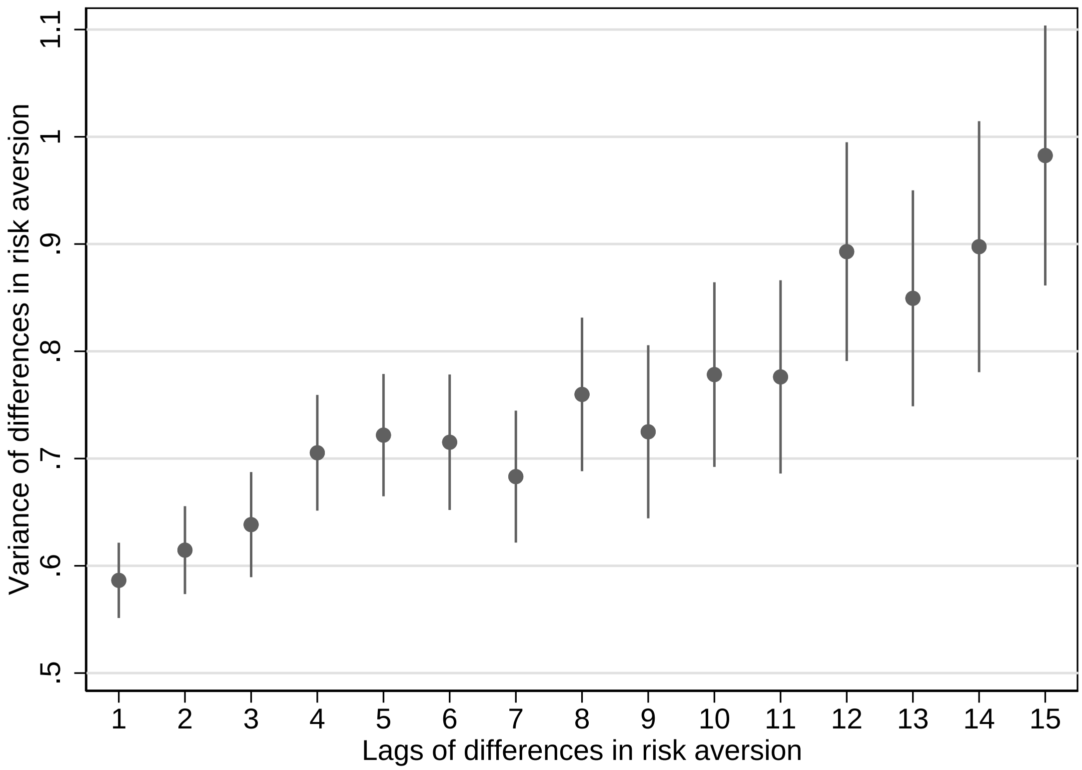
replacement process similar to the derivation of (4), the conditional variance of $P_{it}-P_{i,t-k}$ can be written as:

$$Var[P_{it}-P_{i,t-k}|\bar{X}_{i,t+k}]=(\alpha^k-1)^2\sigma^2_{P*}+\sum_{j=0}^{k}\alpha^{2j}\sigma^2_{\eta}+2\sigma^2_{\varepsilon}, \tag{6}$$

where σ^2_{P*} is the variance of latent preferences. Under the null hypothesis that preferences are time-invariant ($\alpha=1$) and assuming ergodicity of latent preferences, the terms in (6) are monotonically increasing in k at a constant rate. Under the alternative hypothesis that preferences are time-varying ($\alpha<1$) the terms in (6) have a non-linear, potentially non-monotonic relation with k for $k\geq 1$. Put simply, if latent preferences are time-varying, the variance of increasingly lagged differences in measured preferences should have a non-linear, non-monotonic relationship with the lags. If latent preferences are time-stable, these same differences should increase monotonically with the lags.

To perform this test using the DHS data, I construct estimates of the variance of the k-

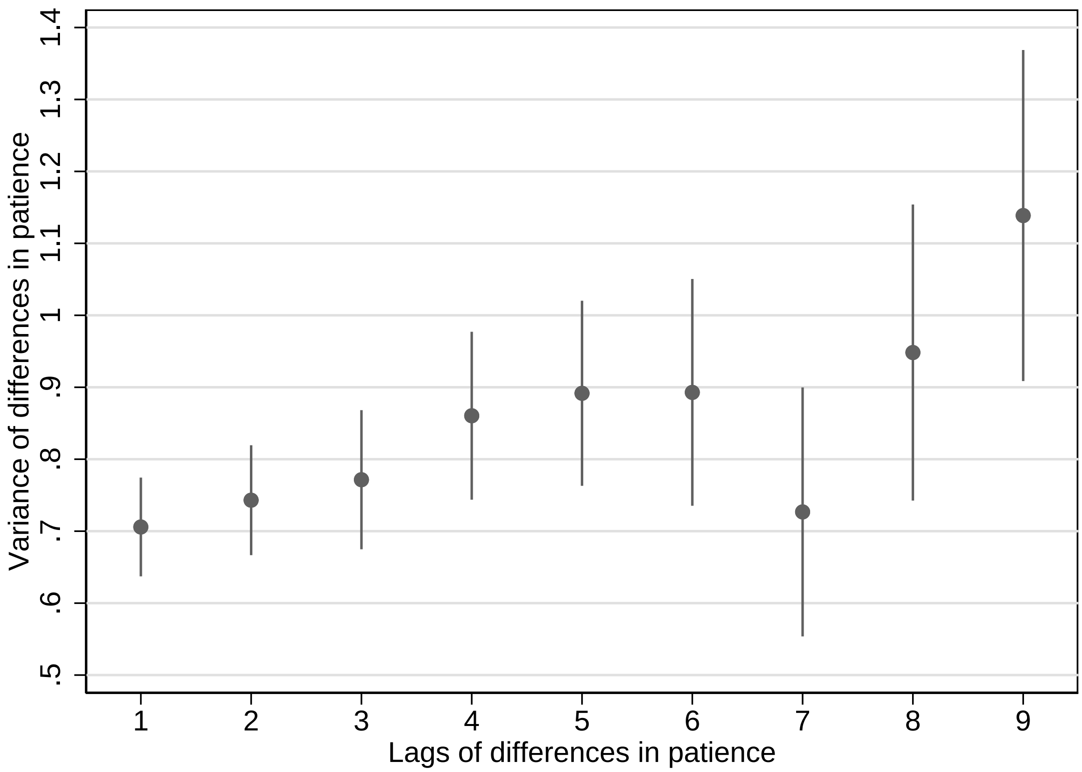
Figure 5: Variance of $k-$lag differences in risk aversion



difference lag in the risk aversion (patience) index for $k=1,...,15$ ($k=1,...,9$). I construct standard errors of for these variance estimates via 499 non-parametric bootstrap replications of this process. Figures 5 and 6 illustrate the results of this process. The figures are consistent with the variance of k–differences in risk aversion and patience increasing with the lags k at a roughly monotonic rate, which would seem to support the idea of time-stable preferences. To formally test monotonic increases at a constat rate, however, I construct a χ^2 test of the null hypothesis of $Var[P_{it}-P_{i,t-1}|\bar{X}_{i,t}]=Var[P_{i,t-1}-P_{i,t-2}|\bar{X}_{i,t-1}]$ for $t=3,...,15$ for risk aversion ($t=3,...,10$ for patience) against the alternative of $Var[P_{it}-P_{i,t-1}|\bar{X}_{i,t}]\neq Var[P_{i,t-1}-P_{i,t-2}|\bar{X}_{i,t-1}]$ for at least some t. The results of this test reject the null of time-invariance of risk aversion ($\chi^2_{13}=33.80$) at all conventional levels, and reject the null of time-invariance of patience ($\chi^2_7=16.58$) at the 5% level.

These variance-based results are consistent with the test results based on the Harris-Tzavalis panel unit root data, and suggest that both risk aversion and patience are time-varying in the DHS. However, the autoregressive parameter estimates in Table 2 still characterize the time variation in preferences, and they indicate that both risk aversion and patience change slowly over time.

Figure 6: Variance of $k-$lag differences in patience



4.3 Identification Alternatives

All the results in Section 3 are calculated by instrumenting lagged preferences with further lags of preferences, they demand measurements of preferences at a minimum of three points in time and these data can be hard to come by. However, any variable Z that fulfills the moment conditions $E[Z \cdot v_{it}] = 0$ can be used to estimate the model's parameters, and this offers a variety of alternatives that are potentially less data-demanding. Below I discuss two such alternatives: splitting multi-item measures of preferences to construct additional instruments, and using heteroscedasticity-based moment conditions for identification.

What if preferences are only measured in two periods? One alternative is to estimate the model's parameters using a split–item instrument. Conceptually, researchers construct indices from multi–item measures because they reduce measurement error compared to single–item measures. Almost always, the underlying assumption behind these indices is that each item measures the underlying construct with some random error. Note, however, that under these assumptions any single item in the index, or any other index constructed as a convex combination of some (or all) of the items will remain a valid measure too. Applying this logic to the setting above, I can construct two indices, instead of just one, out of the multiple items measuring preferences. And I can use one of these indices as an instrument for the other. This approach is similar to the commonly-used technique of using repeated measures of a mismeasured variable to correct for measurement error.[10]

Columns (1) and (2) of Table 4 show the model's parameter estimates, instrumenting a first index of risk aversion and patience (based on items (a), (c), and (f) for risk aversion, and (b), (c), (f), (g), and (j) for patience) with another such index constructed with the remaining items for each construct.[11] This identification method results in similar estimates of the variance of the idiosyncratic and measurement error terms. And, though much closer to one than their OLS counterparts, the method yields markedly lower autoregressive parameter estimates for risk aversion (0.806) and patience (0.729). These discrepancies between these results and those in Table 2 arise likely because the classical measurement error assumption does not hold well for multiple items that were measured at the exact same point in time (i.e., the same day). Probably common factors (e.g., mood or weather) affected respondents' answers to all the preference items the same way. This would create correlated measurement error between the items, and the split-item IV approach would not be able to completely correct for it. This method is likely to yield better results if not all the measurement items are collected at the same time.

What if preferences are measured in only two periods, and only through a single item? In this case one can *still* identify the model's parameters, using heteroscedasticity-based moment conditions. Lewbel (2012) describes this identification, which works as long as $E[X_{it} \cdot \upsilon_{it}^2] \neq 0$, and as long as there is a Z_2 such that $E[Z_2 \cdot \upsilon_{it}] = 0$, where Z_2 can be a subset of X_{it}. The first condition requires some degree of heteroscedasticity in the main equation. The second condition boils down to measurement error being conditionally homoscedastic for at least one observed variable. Lewbel notes that both these conditions often hold in applied microeconomics studies.

Columns (3) and (4) of Table 4 show that estimating the model's coefficients using Lewbel's method yields a slightly larger estimate of the autoregressive parameter for risk aversion, and

[10]Formally, construct two different indices from a set of L different items measuring a preference. The first index, P_{it}^{l}, uses the first l items; the second index, P_{it}^{L-l}, uses the remaining $L-l$ items. If each item's measurement error is classical, the coefficients in Equation (3) can be identified by the moment conditions $E[P_{i,t-1}^{l} \cdot (P_{it}^{L-l} - \beta' X_{it} - \alpha P_{i,t-1}^{L-l})] = 0$.

[11]This way of constructing partitioned indices was based on an a priori intuition of which items were less likely to be correlated through their measurement error, an admittedly subjective exercise. However, with six items measuring risk aversion and 10 items measuring patience, we could construct 52 different index combinations for risk aversion and 21,147 for patience based on the number of possible partitions, or Bell numbers, for both measures. An efficient way to combine these estimates would take their variance–weighted average, yet developing this estimator is beyond the scope if this paper.

a slightly lower one for patience, when compared to estimates in Table 2. Estimates of the variance of idiosyncratic shocks, measurement error and the noise-to-signal ratio are nearly identical for risk aversion, and quite similar for patience. However, I refrain from interpreting these parameters since Lewbel's regressions fail most standard tests for the model's correct identification—a common shortcoming of this method when there are no external instruments.

Table 4:
The dynamic properties of risk aversion and patience under alternative identification strategies

	Split items		Lewbel (2012)	
Preference =	Risk aversion	Patience	Risk aversion	Patience
Lagged preference (α)	0.806	0.729	1.043	0.882
	(0.017)	(0.023)	(0.077)	(0.093)
Idiosyncratic variance (σ_η^2)	0.196	0.226	0.044	0.019
Measurement error variance (σ_ε^2)	0.327	0.343	0.296	0.355
Noise-to-signal ratio (s)	0.486	0.523	0.420	0.550
Controls:	Yes	Yes	Yes	Yes
H_0: Other controls jointly significant	[0.006]	[0.027]	[0.771]	[0.223]

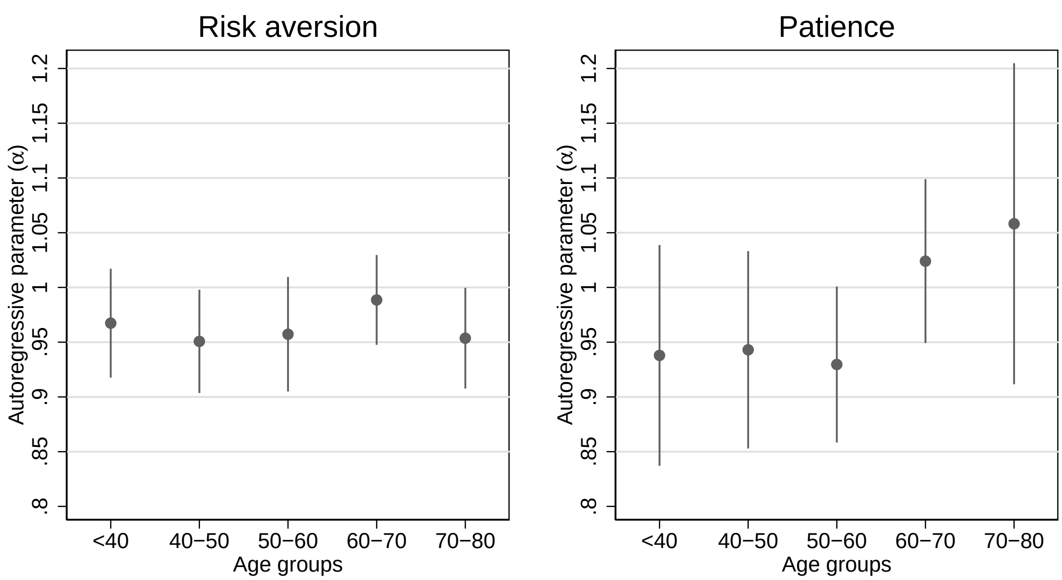
Dynamic properties estimates for risk aversion and patience identified using two alternative methods. Split Sample uses the IV method, instrumenting a first index (based on items (a), (c), and (f) for risk aversion, and (b), (c), (f), (g), and (j) for patience) with another index constructed with the remaining items. Lewbel (2012) uses the heteroscedasticity in other regressors as instruments, and is estimated using the companion Stata package developed by Baum and Schaffer (2015). Controls include income, financial wealth, age, occupation, gender, number of children, and marital status. Robust standard errors clustered at the individual level in parentheses. P-values in square brackets. σ_η^2, σ_ε^2, and s calculated via non-linear least squares using information from 2 lags and netting out X_{it}, as described in Section 2.2.

4.4 Heterogeneous Estimates

The estimates in Section 3 reflect the dynamic parameters for the average respondent, yet they might be hiding relevant sub-group heterogeneity. The most relevant and salient heterogeneity is across age groups. The question of preference stability for different age groups links back to a long-standing literature in psychology claiming that personality becomes stable in early adolescence (Roberts and DelVecchio, 2000; Roberts et al., 2006). Recently age and cohort variation in risk preferences over the life-cycle has been studied in economics by Schurer (2015) and Dohmen et al. (2017). Yet there are other relevant subgroups for which interesting and relevant heterogeneous effects can be estimated. In fact, there are too many of such groups to be canvased in this paper. In this section I only present heterogeneity across a few subgroups in age, occupation, marital status, and gender. These are particularly relevant groups given that many policies selectively target them.

Figures 7 through 10 shows the heterogeneity in $\hat{\alpha}$ in these sub-groups. The estimator $\hat{\alpha}$ remains fairly constant across age groups for risk aversion, yet for patience it is larger for people over 60. Across occupations, $\hat{\alpha}$ remains similar for both preferences, with the exception of patience for disabled people, where it is lower. Divorcees, cohabiting couples, and widows have lower autoregressive parameters in general, but the estimates for these sub-groups are also noisier so it is difficult to draw a clear conclusion. Finally, there seems to be no gender differences in the stability of risk preferences. Women show a higher autoregressive parameter in patience than men, yet these differences are again too noisy to be conclusive.

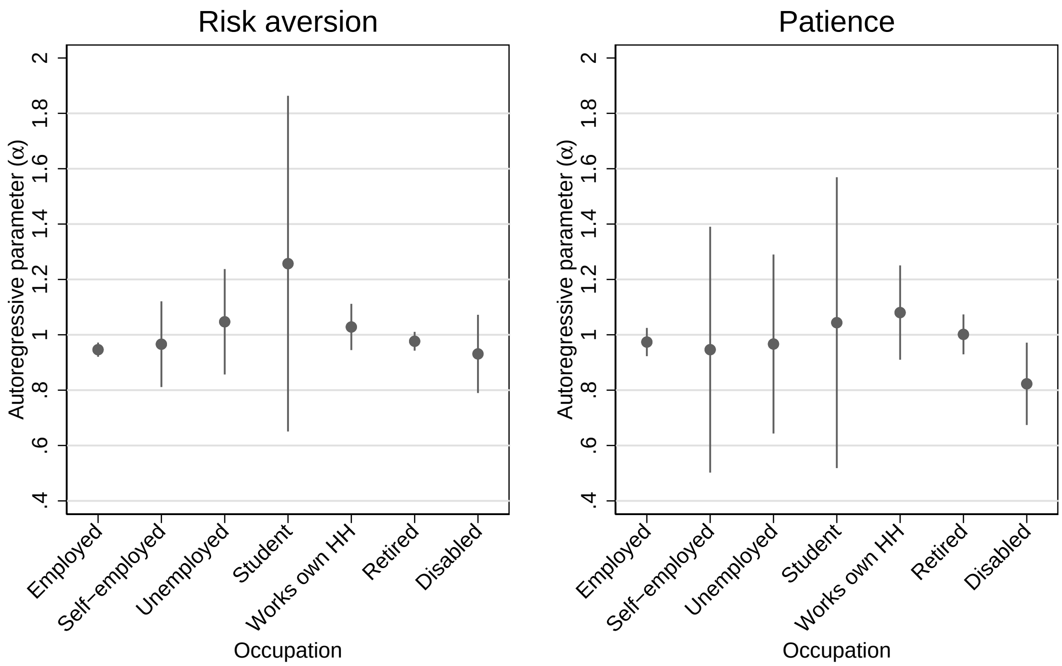
Figure 7: Age heterogeneity in autoregressive parameter (α) estimates



4.5 Stale Preferences and the Severity of the Bias

The evidence in this paper on time-varying risk preferences echoes the concerns of some researchers that inferring risk preferences from "stale" choices can be misleading. This concern is explicitly expressed in Einav et al. (2012) but applies to the studies that elicit risk preferences using either (hypothetical) lottery choices (e.g., Barsky et al., 1997; Andersen et al., 2008; Dohmen

Figure 8: Occupation heterogeneity in autoregressive parameter (α) estimates

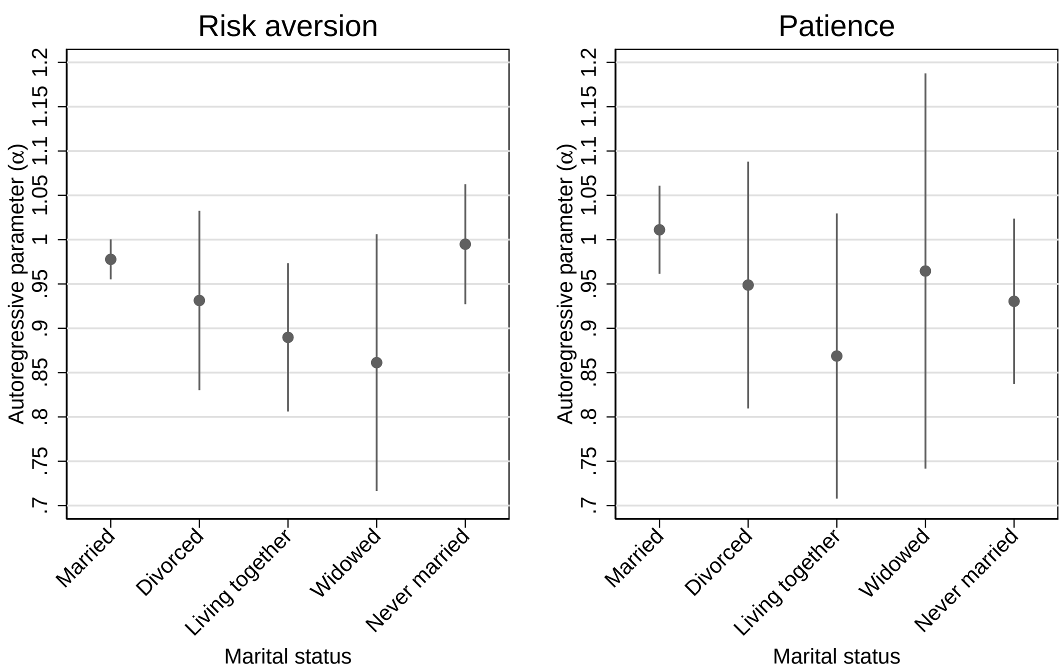


et al., 2011) or observed choices (e.g., Cutler and Glaeser, 2005; Einav et al., 2012). If risk preferences are stable, one could relate risk preferences measured at one point in time to outcomes observed much earlier or later. However, since risk preferences do change, extrapolating preferences can often lead to underestimating the effect of preferences on behavior. I argue that conclusions can change substantially when using stale preferences through an example using risk aversion.

To estimate the severity of the bias from extrapolation of preferences, I simulate a contemporaneous relation between an outcome y_t (say, portfolio risk-taking) and risk aversion r_t (measured without error, since it is not an important component for this exercise), and then I check what would happen if I use increasingly stale measures of the risk aversion (e.g., r_{t-1},r_{t-2},...) to predict the outcome. I simulate risk preferences in the data generating process (DGP) evolving as in equations (2) and (3) and postulate a one-to-one contemporaneous relation between y_t and r_t (i.e., a true coefficient of $\gamma=1$ when y_t is regressed on r_t).[12] The simulation was done 1,000 times for N=3,000 individuals over K=10 to mimic the DHS data.

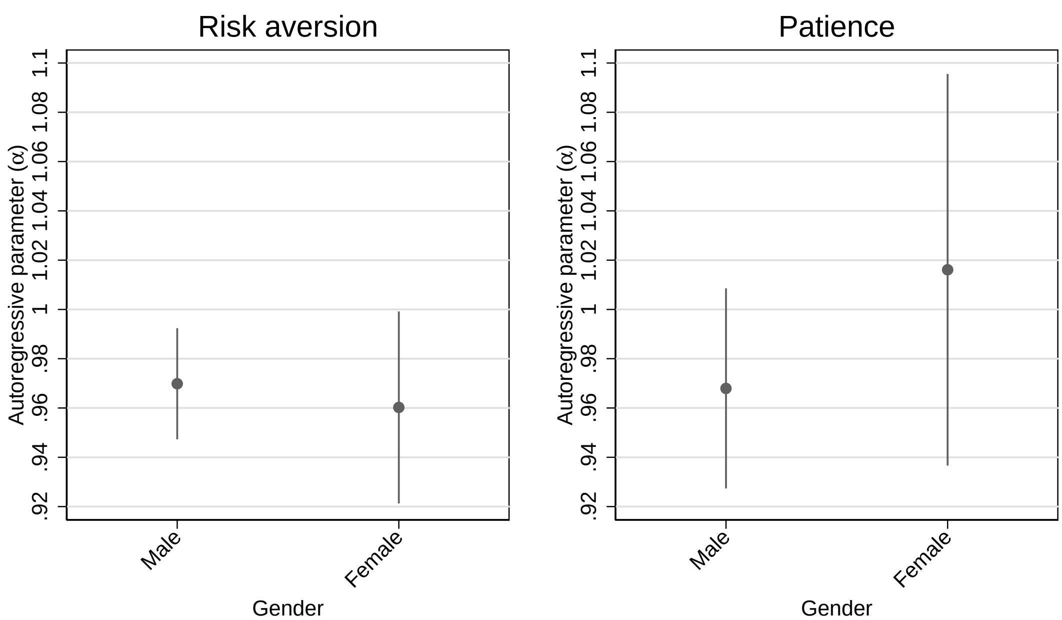
[12]The DGP is: $y_t=\gamma r_t+e_t$ with $r_t=r_t^*+\eta_t$ and $r_t^*=\alpha r_{t-1}^*+\varepsilon_t$. The parameters in the simulation were calibrated based on Column (4) of Table 2. $e_t\sim\mathcal{N}(0,1)$.

Figure 9: Marital status heterogeneity in autoregressive parameter (α) estimates



The exercise shows that, as expected, when regressing the outcome on the contemporaneous risk preferences the true coefficient estimate is well-identified and precisely estimated. As increasingly stale measures of risk preferences are used, however, the estimates drift farther apart from the true effect. The discrepancies between the true effect and the estimates is most severe when estimating the effect via OLS since the estimates are biased because of the staleness of risk preferences and the idiosyncratic shocks to preferences, which act as measurement error in this scenario. Instrumenting current preferences with lagged preferences decrease the bias considerably. Yet, as the risk preference become increasingly stale, the estimates also become severely biased. In this simple exercise, tailored to resemble the DHS data, the bias introduced by the use of stale preferences would lead us to reject the true null of $\gamma = 1$ when using preferences measured over more than 4 years. The conclusion is that assuming preferences are stable can lead to the wrong conclusions, even when preferences evolve slowly and are well measured.

Figure 10: Gender heterogeneity in autoregressive parameter (α) estimates



5 Conclusions

This paper provides a first formal statistical test for the stability of preferences. It also provides metrics for different important features of the dynamics of preferences that should be considered separately, including their predictable change by observable characteristics, the variance of idiosyncratic shocks to them, and their measurement error. And, perhaps most importantly for applied economists, its method is simple to implement and the paper comes with its own Stata code, `dynreg`, which I am happy to provide.

To illustrate the use of this method, I estimate the dynamic properties of risk aversion and patience in long-running panel data. I find that patience is time-stable but risk aversion is not. However, both risk aversion and patience change very slowly over time. This slow rate of change means that applied researchers interested in these preferences can essentially consider them as time-invariant over the span of a few years. I also show that the variance of persistent shocks to both preferences are tiny; that levels, but not changes, in preferences are well predicted by observable characteristics; and that over 30 percent of the variation in both preference measures is measurement error. Finally, in a number of extensions, I describe a second test for the stability of preferences that relies on different identifying variation, and I discuss alternative identification for the method's parameters.

This method, like most other methods that provide precise answers to questions, relies on some assumptions. Some of those assumptions are innocuous and just serve to simply matter, and a few of them are crucial for the method to work. The important takeaway, however, is that in this paper I provide the very first way to formally test one of the core assumptions in economics, and that this test is simple enough to be practicable by virtually all researchers. This alone should be enough to stop blindly relying on the assumption of preference stability, and start testing it.

References

ANDERSEN, S., G. W. HARRISON, M. I. LAU, AND E. E. RUTSTRÖM (2008): "Eliciting risk and time preferences," *Econometrica*, 76, 583–618.

ARELLANO, M. AND S. BOND (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic Studies*, 58, 277–297.

BARSKY, R., F. JUSTER, M. KIMBALL, AND M. SHAPIRO (1997): "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics*, 112, 537–579.

BAUM, C. F. AND M. E. SCHAFFER (2015): "IVREG2H: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments," *Statistical Software Components*.

BORGHANS, L. AND B. GOLSTEYN (2006): "Time discounting and the body mass index: Evidence from the Netherlands," *Economics & Human Biology*, 4, 39–61.

BUCCIOL, A. AND R. MINIACI (2018): "Financial Risk Propensity, Business Cycles and Perceived Risk Exposure," *Oxford Bulletin of Economics and Statistics*.

CAMERON, A. C. AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*, Cambridge university press.

CARROLL, C. AND A. SAMWICK (1997): "The nature of precautionary wealth," *Journal of Monetary Economics*, 40, 41–71.

COBB-CLARK, D. A. AND S. SCHURER (2013): "Two Economists' Musings on the Stability of Locus of Control," *The Economic Journal*, 123, F358–F400.

CUTLER, D. M. AND E. GLAESER (2005): "What Explains Differences in Smoking, Drinking, and Other Health-Related Behaviors?" *The American Economic Review*, 95, 238–242.

DOHMEN, T., A. FALK, B. H. GOLSTEYN, D. HUFFMAN, AND U. SUNDE (2017): "Risk attitudes across the life course," *The Economic Journal*, 127, F95–F116.

DOHMEN, T., A. FALK, D. HUFFMAN, U. SUNDE, J. SCHUPP, AND G. WAGNER (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*.

DOHMEN, T., H. LEHMANN, AND N. PIGNATTI (2016): "Time-varying individual risk attitudes over the Great Recession: A comparison of Germany and Ukraine," *Journal of Comparative Economics*, 44, 182–200.

DUERSCH, P., D. ROMER, AND B. ROTH (2017): "Intertemporal stability of uncertainty preferences," *Journal of Economic Psychology*, 60, 7 – 20, stability of Preferences.

EINAV, L., A. FINKELSTEIN, I. PASCU, AND M. R. CULLEN (2012): "How General Are Risk Preferences? Choices under Uncertainty in Different Domains," *The American Economic Review*, 102, 2606–2638.

GOLSTEYN, B. AND H. SCHILDBERG-HORISCH (2017): "Challenges in research on preferences and personality traits: Measurement, stability, and inference," *Journal of Economic Psychology*, 60, 1 – 6, stability of Preferences.

GUISO, L., P. SAPIENZA, AND L. ZINGALES (2018): "Time varying risk aversion," *Journal of Financial Economics*.

HANAOKA, C., H. SHIGEOKA, AND Y. WATANABE (forthcoming): "Do Risk Preferences Change? Evidence from the Great East Japan Earthquake," *American Economic Journal: Applied Economics*.

HARDARDOTTIR, H. (2017): "Long term stability of time preferences and the role of the macroeconomic situation," *Journal of Economic Psychology*, 60, 21 – 36, stability of Preferences.

HARRIS, R. D. AND E. TZAVALIS (1999): "Inference for unit roots in dynamic panels where the time dimension is fixed," *Journal of Econometrics*, 91, 201–226.

HOFFMANN, A. O., T. POST, AND J. M. PENNINGS (2013): "Individual investor perceptions and behavior during the financial crisis," *The Journal of Banking & Finance*, 37, 60–74.

KAPTEYN, A. AND F. TEPPA (2011): "Subjective measures of risk aversion, fixed costs, and portfolio choice," *Journal of Economic Psychology*, 32, 564–580.

KRUPKA, E. L. AND M. STEPHENS (2013): "The stability of measured time preferences," *Journal of Economic Behavior & Organization*, 85, 11–19.

LEVIN, A., C.-F. LIN, AND C.-S. J. CHU (2002): "Unit root tests in panel data: asymptotic and finite-sample properties," *Journal of Econometrics*, 108, 1–24.

Lewbel, A. (2012): "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models," *Journal of Business & Economic Statistics*, 30.

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

Meier, S. and C. D. Sprenger (2015): "Temporal stability of time preferences," *Review of Economics and Statistics*, 97, 273–286.

Mundlak, Y. (1978): "On the pooling of time series and cross section data," *Econometrica*, 46, 69–85.

Roberts, B. and W. DelVecchio (2000): "The rank-order consistency of personality traits from childhood to old age: a quantitative review of longitudinal studies." *Psychological bulletin*, 126, 3.

Roberts, B., K. Walton, and W. Viechtbauer (2006): "Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies." *Psychological Bulletin*, 132, 1–25.

Sahm, C. R. (2012): "How much does risk tolerance change?" *The Quarterly Journal of Finance*, 2.

Salamanca, N., A. de Grip, D. Fouarge, and M. R. (2016): "Locus of control and investment in risky assets," *IZA Discussion Paper Series*, No. 10407.

Schurer, S. (2015): "Lifecycle patterns in the socioeconomic gradient of risk preferences," *Journal of Economic Behavior & Organization*, 119, 482–495.

Valente, C. and S. Sokullu (2018): "Individual Consumption in Collective Households: Identification Using Panel Data with an Application to PROGRESA," *Unpublished mimeo*.

Warneryd, K. (1996): "Risk attitudes and risky behavior," *Journal of Economic Psychology*, 17, 749–770.

Appendix

Table A.1: Summary statistics for control variables in the risk aversion and patience samples

Item	Mean	S.D.	Min	Max
Risk aversion ($N=10{,}826$)				
Gross household income (€1,000)	46.07	23.73	1	148
Net household wealth (€1,000)	74.66	113.92	1	995
Age	54.48	14.02	18	80
Number of children in household	0.65	1.03	0	6
Female	0.39	0.49	0	1
Occupation:				
Employed	0.51	0.50	0	1
Self-employed	0.03	0.16	0	1
Unemployed	0.02	0.13	0	1
Student	0.01	0.11	0	1
own household	0.11	0.31	0	1
Retired	0.28	0.45	0	1
Disabled	0.04	0.19	0	1
Marital status:				
Married	0.71	0.45	0	1
Divorced	0.04	0.20	0	1
Living together	0.07	0.25	0	1
Widowed	0.03	0.17	0	1
Never married	0.11	0.32	0	1
Missing	0.04	0.19	0	1
Patience ($N=3{,}710$)				
Gross household income (€1,000)	42.53	21.40	1	147
Net household wealth (€1,000)	71.03	116.18	1	1000
Age	52.47	13.82	18	80
Number of children in household	0.75	1.09	0	6
Female	0.40	0.49	0	1
Occupation:				
Employed	0.53	0.50	0	1
Self-employed	0.02	0.15	0	1
Unemployed	0.02	0.13	0	1
Student	0.01	0.12	0	1
own household	0.14	0.34	0	1
Retired	0.24	0.43	0	1
Disabled	0.04	0.19	0	1
Marital status:				
Married	0.71	0.45	0	1
Divorced	0.05	0.21	0	1
Living together	0.06	0.23	0	1
Widowed	0.02	0.13	0	1
Never married	0.12	0.33	0	1
Missing	0.04	0.21	0	1

Summary statistics for controls in the risk aversion and patience main samples.