Estimating Individual Responses when Tomorrow Matters*

Stéphane Bonhomme
University of Chicago

Angela Denis
Banco de España

June 8, 2023

Abstract

We propose a regression-based method to account for the role of expectations in how individuals respond to a policy that changes the economic environment. We provide conditions under which average partial effects based on regression estimates recover structural counterfactual effects. For estimation, we rely on subjective beliefs data and propose a practical three-step method. We illustrate our approach in a model of consumption and saving, and focus on the impact of an income tax that not only changes current income but also affects beliefs about future income. In an application to Italian survey data, we find that accounting for beliefs matters to assess how tax policies impact consumption decisions.

JEL codes: C10. C50.

Keywords: Dynamics, subjective expectations, beliefs, semi-structural estimation.

*We thank Steven Durlauf, Lars Hansen, Jim Heckman, Philip Heiler, Michael Keane, Pierre-Carl Michaud, Claudio Michelacci, Eduardo Morales, Chris Taber, Wilbert Van der Klaauw, Daniel Wilhelm, and participants at various places for comments. Contact: sbonhomme@uchicago.edu; angela.denis@bde.es. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Banco de España or the Eurosystem.

1 Introduction

Economists often seek to assess how changes in the economic environment affect individual decisions. A leading example is the *ex ante* evaluation of policies that have not yet taken place. However, a key challenge is that, when the environment changes, individual decision rules are generally affected as well. In dynamic settings with uncertainty, one not only needs to account for the contemporaneous effects of the change, but also for its impact through a change in expectations.

A common approach in empirical work is to regress the outcome on the covariate that one is interested in shifting in the counterfactual (e.g., under a new policy), and to compute an average partial effect based on this regression. Average partial effects can be structurally interpreted as counterfactual policy effects, in a framework where the covariate is a state variable in the individual decision problem, and the regression function is interpreted as the agent's decision rule (Stock, 1989). However, underlying this interpretation is the assumption that the decision rule is invariant to the change of interest. This invariance assumption can be restrictive in many settings where agents' beliefs about the future matter.

Consider the introduction of a permanent income tax in a standard model of consumption and saving (see Deaton, 1992, for a textbook treatment). The effect of the tax can be estimated by regressing consumption and income, and computing an average partial effect associated with the tax change. However, such an effect is likely to be empirically misleading, since both current income and beliefs about future income will be affected by the tax. Not accounting for the change in beliefs will produce biased predictions of the effect of the tax, as shown by Lucas (1976) in his influential critique.

As a second example, consider the effect of a change in the weather process in a model of agricultural production where farmers choose dynamic inputs (such as irrigation or a fertilizer) based on their forecasts of future weather. In addition to affecting contemporaneous weather conditions, a change in the weather process will affect farmers' beliefs about future weather, which may lead them to modify their input choices. Not accounting for farmers' adaptation will bias calculations of the impact of a change in the weather process (Deschênes and Greenstone, 2007, Burke and Emerick, 2016).

In this paper, our aim is to study and estimate average partial effects in an analysis that explicitly accounts for the role of individual expectations. In our setup, individual beliefs are determinants of decisions, and they enter as additional state variables in the agent's decision problem. We show how to assess the total effect of a policy or other counterfactual change, and how to decompose it into a contemporaneous effect, where beliefs are held fixed, and a purely dynamic effect that reflects the change in beliefs.

To implement this approach we rely on data on subjective expectations. Beliefs data are increasingly available in a variety of settings (Manski, 2004). We propose to construct estimates of subjective distributions based on such survey responses, and to account for those in the definition and estimation of average partial effects. There are many examples of the use of beliefs data on the right-hand side of a regression.[1] Here we show how to interpret the estimates of such regressions, and we provide conditions under which those can be used for counterfactual prediction.

To interpret regression-based average partial effects, we propose a structural dynamic framework where agents choose actions based on their beliefs about the future. In the spirit of semi-structural approaches, we use the framework to justify the use of average partial effects, but we do not need to specify or estimate a structural model.[2] As a result, the counterfactuals we focus on are restricted to changes in the process of covariates and beliefs, and our approach cannot answer counterfactual questions related to changes in preferences or technology, for example.[3]

In the structural framework, beliefs are time-varying state variables in the agent's decision problem. Variation in beliefs over time is crucial, since it allows us to control for preference heterogeneity, which we assume to be constant over time. We assume that beliefs are exogenous, and that, together with other exogenous state variables, they follow a first-order Markov process. While belief exogeneity is a central feature of our framework, we show the Markovian assumption is compatible with various popular models of belief formation.

In the framework the agent's decision rule is a function of exogenous state variables (such as income or the weather), including the beliefs, as well as endogenous dynamic state variables

[1]For example, Guiso and Parigi (1999) study how a firm's investment depends on its beliefs about future demand, Hurd, Smith, and Zissimopoulos (2004) study the effects of subjective survival probabilities on decisions about retirement and social security claims, Dominitz and Manski (2007) analyze how beliefs about equity returns affect portfolio choice, Bover (2015) studies how subjective expectations about future home prices affect car and secondary home purchases, and Attanasio, Cunha, and Jervis (2019) study how parental investment in children is influenced by beliefs about the production function.

[2]Our focus on the estimation of policy effects without a full structural model follows the spirit of Marschak (1953), Ichimura and Taber (2000, 2002), and Keane and Wolpin (2002a,b), among others; see also Wolpin (2013).

[3]There is a growing literature on the combination of structural models and subjective beliefs data, see among others Van der Klaauw and Wolpin (2008), Delavande (2008), Van der Klaauw (2012), Stinebrickner and Stinebrickner (2014), Arcidiacono, Hotz, Maurel, and Romano (2020), Wiswall and Zafar (2015), and Koşar and O'Dea (2022); see also the recently released handbook on economic expectations (Bachmann, Topa, and van der Klaauw, 2022).

(such as assets or capital). We can assess the effects of an exogenous change in those arguments by computing average partial effects which, unlike in the static case, account for changes in beliefs. Such effects correspond to well-defined structural counterfactuals under the assumption that the dynamic decision rule is invariant to the change. Hence, we rely on a weaker type of invariance assumption than standard average partial effects that do not allow beliefs to respond to changes, yet a certain form of invariance is still needed to structurally interpret average partial effects in our setup.

In our first example, the consumption function depends on current income, assets, and time-varying beliefs about next period's income. To structurally interpret an average partial effect as the effect of a counterfactual tax, one needs to assume that preferences, discounting, and assets returns are not affected by the tax, and that the belief updating rule, which governs the dynamic process of beliefs, remains constant as well. Hence, while we allow changes in the environment to modify beliefs about the near future, the long-run process of those beliefs remains invariant to the change. We show this assumption is not needed in a permanent-transitory version of the consumption model for a proportional income tax, and in an extension we show how to relax it when belief data over longer horizons are available. To illustrate this example, we estimate average partial effects in data simulated from a structural consumption model, using two different processes of individual income expectations, and we show they correctly recover the model-based counterfactual impacts.

Our empirical analysis of average partial effects in dynamic settings proceeds in three easy-to-implement steps. In the first step, we estimate the belief distributions. To account for the fact that survey responses on subjective beliefs tend to be coarse, we assume that belief distributions depend on a finite-dimensional parameter. However, we also describe semi- and nonparametric extensions that can be implemented with rich beliefs data. In the second step we estimate the regression function (i.e., the agent's decision rule). In the third step, we use these estimates to compute the impact of a change in its arguments (i.e., in the state variables). Without additional modeling assumptions, nonparametric identification is limited to the empirical support of the covariates, including beliefs.

As an empirical illustration, we study how consumption decisions depend on current income and beliefs about future income. We rely on Italian data from the Survey on Household Income and Wealth (SHIW), which contains information on respondents' probabilistic income expectations.[4] We then use our approach to predict the impact of various counterfactual income taxes,

[4]Elicited beliefs about future income are increasingly available. For example, Pistaferri (2001), Guiso, Jappelli, and Pistaferri (2002), and Kaufmann and Pistaferri (2009) use data on income expectations in the SHIW

involving transitory or permanent increases in marginal tax rates, and a change in the degree of progressivity of the tax. We find that, conditional on current income, income beliefs shape consumption responses, and that they matter for predicting the effects of income taxes.

The outline is as follows. In Section 2 we introduce average partial effects for dynamic settings. In Section 3 we describe a structural framework and discuss the interpretation of average partial effects in this context. We present two examples in Section 4. We study identification and estimation in Section 5, and we present our consumption application in Section 6. Finally, in Section 7 we describe some extensions of the approach.

2 Average partial effects when tomorrow matters

Consider an individual outcome y_{it} that depends on some covariates x_{it} and z_{it}. Suppose that, for some function g_i,

$$y_{it} = g_i(x_{it}, z_{it}) + \varepsilon_{it}, \tag{1}$$

where ε_{it} has zero mean given x_{it} and z_{it}. To fix ideas, we will refer to the case where y_{it} is a measure of consumption, x_{it} is income, and z_{it} includes other determinants such as assets.

Consider an exogenous change in x_{it}, from $x_{it} = x$ to some other value $x_{it} = x^{(\delta)}$. For example, if one is interested in a mean shift of (log) income by a δ amount, corresponding to a proportional tax or subsidy, one will set $x^{(\delta)} = x + \delta$. A standard average partial effect associated with the change in x_{it} is then

$$\Delta_i^{\text{APE}}(\delta, x, z) = g_i(x^{(\delta)}, z) - g_i(x, z),$$

possibly averaged across individual observations. By estimating quantities such as Δ_i^{APE}, one can document how individual responses vary across individuals and across values of x and z.

However, to interpret Δ_i^{APE} as the impact on outcomes when x_{it} changes from x to $x^{(\delta)}$, one needs to assume that, as x_{it} changes while z_{it} is kept constant, the function g_i remains constant (Stock, 1989). This invariance assumption is often implausible in applications where dynamics matter. Indeed, in many settings where the current value of x_{it} changes, beliefs about future x_{it}'s, which are implicitly contained in the function g_i, are likely to change as well. For example, under a (permanent) tax, both current income and beliefs about future income change.

in combination with models of consumption and saving. Stoltenberg and Uhlendorff (2022) estimate a structural model with subjective income expectations using the same data. Attanasio, Kovacs, and Molnar (2020) combine data on subjective expectations with data on actual income and estimate an Euler equation for consumption.

Our approach to alleviate this well-known issue is to augment (1) by including beliefs about future x_{it} values as additional determinants of y_{it}. Letting π_{it} denote the subjective distribution of $x_{i,t+1}$ at time t, we postulate that, for some function ϕ_i,

$$y_{it} = \phi_i(x_{it}, \pi_{it}, z_{it}) + \varepsilon_{it}, \tag{2}$$

where ε_{it} has zero mean given x_{it}, π_{it} and z_{it}. In the consumption example, this amounts to including income beliefs as additional determinants of the consumption function.

In model (2), we will be interested in documenting the effects of a change from $x_{it} = x$ to $x_{it} = x^{(\delta)}$, associated with a change in beliefs from $\pi_{it} = \pi$ to $\pi_{it} = \pi^{(\delta)}$. As an example, consider again the effect of shifting current (log) income x_{it} by a δ amount. In this case, $\pi^{(\delta)}$ is the belief about future income $x_{i,t+1}$ under the δ mean shift. Such a joint change has two distinct effects on outcomes: a contemporaneous effect associated with the change in x_{it}, and a dynamic effect associated with the change in beliefs π_{it}.

In this setup, we define the total average partial effect, or TAPE, as

$$\Delta_i^{\text{TAPE}}(\delta, x, \pi, z) = \phi_i(x^{(\delta)}, \pi^{(\delta)}, z) - \phi_i(x, \pi, z). \tag{3}$$

We then further decompose the total average partial effects as the sum of two terms: a contemporaneous APE (or CAPE), where beliefs are held constant, and a dynamic APE (or DAPE), which solely captures the change in beliefs. Formally, we decompose

$$\Delta_i^{\text{TAPE}}(\delta, x, \pi, z) = \underbrace{\phi_i(x^{(\delta)}, \pi, z) - \phi_i(x, \pi, z)}_{=\Delta_i^{\text{CAPE}}(\delta, x, \pi, z)} + \underbrace{\phi_i(x^{(\delta)}, \pi^{(\delta)}, z) - \phi_i(x^{(\delta)}, \pi, z)}_{=\Delta_i^{\text{DAPE}}(\delta, x, \pi, z)}. \tag{4}$$

To interpret Δ_i^{TAPE} as the impact on outcomes when x_{it} changes from x to $x^{(\delta)}$ and π_{it} changes from π to $\pi^{(\delta)}$, one needs to assume that the function ϕ_i remains invariant in the counterfactual. Although this assumption is not without loss of generality (and we will discuss it in the context of a structural framework in the next section), it is weaker than the assumption that g_i in (1) is invariant to the change. The key difference is that, unlike (1), (2) explicitly accounts for variation in beliefs.

In the next section we will describe a class of structural models under which (2) is the individual optimal decision rule in the economic problem. This will allow us to transparently discuss the assumptions needed to structurally interpret the above average partial effects (TAPE, CAPE and DAPE).

The structural framework has two main features. First, x_{it} and π_{it} jointly follow an exogenous first-order Markov process. This implies that x_{it}, π_{it}, and z_{it} are the state variables of the economic problem (in addition to some shocks subsumed in ε_{it}). Belief exogeneity is a key feature of the framework. In turn, the first-order Markov assumption imposes restrictions on the belief formation process, however we show it is satisfied in several popular models of beliefs. Second, in the structural model, ϕ_i depends on preferences, discounting, the law of motion of z_{it}, and the law of motion of the beliefs π_{it}. To guarantee the invariance of ϕ_i, we will assume that none of these quantities varies under the policy change.

Assuming that the law of motion of the beliefs, which we denote as ρ_i, is invariant requires that, while agents account for the impact of the change on their beliefs about $x_{i,t+1}$, the way they update their beliefs after period $t+1$ is unaffected. Under this assumption, ρ_i is an individual "type" that is invariant to the change. We will see that this assumption is satisfied in a popular version of the consumption example.[5]

Finally, note that, when beliefs matter in (2), an approach based on (1) is incorrect for two reasons. The first one is that beliefs π_{it}, which are generally correlated with x_{it}, are omitted variables in (1). Hence, not controlling for π_{it} gives incorrect estimates of the contemporaneous APE. The second one is that relying on (1) makes it impossible to estimate the total APE, and to decompose it into contemporaneous and dynamic average partial effects. Hence, when (2) holds, Δ_i^{APE} is not economically interpretable in general.

3 Structural interpretation

In this section we describe a structural dynamic framework where individual decision rules take the form (2), and we provide a structural interpretation for average partial effects.

3.1 Economic environment

Consider an individual i's intertemporal decision making process in discrete time. In the presentation we focus on a stationary infinite-horizon environment. However, in Remark 1 we will show how to apply the framework to finite-horizon environments.

In every period t, the individual observes four vectors of state variables: an exogenous state x_{it}, another exogenous state π_{it} that represents agents' beliefs about future values $x_{i,t+1}$, an endogenous state z_{it}, and a vector of taste shifters ν_{it}, or "shocks", which are i.i.d. over time.

[5] Relaxing this assumption is conceptually straightforward in our framework, by defining π_{it} in (2) as beliefs about a sequence of future x's, $x_{i,t+1}, x_{i,t+2}, ..., x_{i,t+S}$. However, doing so imposes stronger demands on the data. We will return to this point in the extensions section.

Given the state variables, the individual chooses an action y_{it}. We specify the evolution of state variables (as perceived by the agent) as follows.

Assumption 1. *(State variables)*

1. *[Exogenous state variables]* (x_{it}, π_{it})*, t=1,2,...., is first-order Markov.*
2. *[Beliefs] For all* x, π*,*
$$(x_{i,t+1} \,|\, x_{it} = x, \pi_{it} = \pi) \sim \pi.$$
3. *[Belief updating] For all* $\widetilde{x}, \pi, x$*,*
$$(\pi_{i,t+1} \,|\, x_{i,t+1} = \widetilde{x}, \pi_{it} = \pi, x_{it} = x) \sim \rho_i(\cdot; \widetilde{x}, \pi, x).$$
4. *[Endogenous state variables] There exists a non-stochastic function* γ_i *such that*
$$z_{i,t+1} = \gamma_i(z_{it}, x_{it}, y_{it}).$$
5. *[Shocks]* $\nu_{it} \sim \tau_i$*, i.i.d. over time and independent of all* x_{it}*'s and* π_{it}*'s.*

Together, parts 1, 2 and 3 in Assumption 1 specify the law of motion of states (x_{it}, π_{it}) as perceived by the agent: part 1 specifies the exogenous state (x_{it}, π_{it}) as first-order Markov, part 2 defines π_{it} to be the individual's beliefs about future values $x_{i,t+1}$, and part 3 specifies the updating rule for π_{it}. Note that we do not impose a rational expectations assumption, so perceived and realized laws of motion may not coincide. This is an important point since we will show that, when subjective expectations data are available, it is not necessary to impose rational expectations in order to estimate decision rules.

It is important to note that in our setup beliefs are exogenous, in the sense that they are not affected by the agent's past actions. This can be restrictive in situations where individual beliefs are shaped by past choices. For example, in Lochner (2007), beliefs about the probability of arrest depend on past decisions to commit crime and whether those decisions led to an arrest. This type of belief endogeneity is not covered by our framework.

Part 4 in Assumption 1 specifies the endogenous state z_{it}, allowing for feedback from actions y_{it} to future states $z_{i,t+1}$.[6] While such feedback is not allowed for the exogenous states (x_{it}, π_{it}), we outline a generalization where agents have state-contingent beliefs in the extensions section. Lastly, part 5 introduces taste shifters ν_{it}, which enter utility and will be subsumed in the error term ε_{it} in (2).

[6] The setup is unchanged if, in addition to (z_{it}, x_{it}, y_{it}), $z_{i,t+1}$ also depends on beliefs π_{it}.

3.2 Compatibility with belief formation models

Under Assumption 1, beliefs follow a first-order Markov process jointly with the other exogenous state variables. We now illustrate that this assumption is consistent with several models of belief formation in economics, see Pesaran and Weale (2006) for references.

As a first example, suppose that agents have rational expectations, and that agent i's information set at time t is $\Omega_{it} = \{x_i^t, \eta_i^t\}$,[7] where $x_{it} = \eta_{it} + \varepsilon_{it}$, η_{it} is homogeneous first-order Markov, and ε_{it} is independent of η_{it} with a stationary distribution. An example is a permanent-transitory specification of the income process, which we will study in our consumption example. In this setup, π_{it} is the distribution of $x_{i,t+1}$ given Ω_{it}, which is also the distribution of $x_{i,t+1}$ given η_{it}. Given that η_{it} is first-order Markov, part 3 in Assumption 1 is thus satisfied.[8] However, it generally fails if η_{it} is not first-order Markov, or if individuals have additional signals about future $\eta_{i,t+s}$ and $\varepsilon_{i,t+s}$.

As a second example, suppose that $x_{it} = \alpha_i + \varepsilon_{it}$, yet agents do not know α_i and try to learn it given the observations from x_{it}. We show in Appendix A that, when ε_{it} is Gaussian and Bayesian agents have Gaussian priors about α_i and rational expectations, x_{it} and the beliefs π_{it} jointly follow a first-order Markov process, consistently with Assumption 1. Hence, our setup can allow for some form of learning. However, it does not allow for learning from past choices, since we assume that beliefs are exogenous, nor does it allow for agents to learn from other agents.

Our setup is also compatible with some models of non-rational expectations. As a third example, consider a simple model of adaptive expectations, where mean beliefs evolve as

$$\mathbb{E}_{\pi_{it}}(x_{i,t+1}) = \mathbb{E}_{\pi_{i,t-1}}(x_{it}) + \lambda_i \left(x_{it} - \mathbb{E}_{\pi_{i,t-1}}(x_{it})\right). \tag{5}$$

In the terminology of Armona, Fuster, and Zafar (2019), individuals with $\lambda_i > 0$ are "extrapolators", those with $\lambda_i = 0$ are "non-updators", and those $\lambda_i < 0$ are "mean reverters". Assumption 1 is satisfied if (5) holds, and, say, beliefs are normal with constant variance σ_i^2. More generally, Assumption 1 is consistent with adaptive expectations models where the entire belief distribution π_{it} depends on $\pi_{i,t-1}$ and x_{it}.

Our framework relies on the assumption that, given x_{it} and z_{it}, π_{it} is sufficient for the history of past x, z and π variables. This sufficiency assumption holds in the above examples. However, our assumption that $(x_{it}, z_{it}, \pi_{it})$ are the relevant state variables implicitly restricts

[7] Here we denote as $w^t = (w_t, w_{t-1}, w_{t-2}, ...)$ the history of the random variable w_t.

[8] More generally, it suffices that (x_{it}, η_{it}) be first-order Markov and that the mapping $(x_{it}, \eta_{it}) \mapsto (x_{it}, \pi_{it})$ be injective.

the information that is relevant to individual decisions. An advantage of our approach is that, since beliefs π_{it} are state variables, we can perform counterfactual exercises that account for changes in beliefs without the need for a full-fledged structural model.

3.3 Decisions and policy rule

Let $u_i(y_{it}, x_{it}, z_{it}, \nu_{it})$ denote period t's contemporaneous payoffs. Here the action may be continuous or discrete, so our framework covers structural dynamic discrete choice models as well as models with continuous choices.[9] Let β_i denote the time discount factor. The individual solves the infinite horizon program

$$(y_{i,1}, y_{i,2}, ...) = \max_{(a_1,a_2,...)} \mathbb{E}\left[\sum_{t=1}^{\infty} \beta_i^{t-1} u_i\left(a_t, x_{it}, z_{it}, \nu_{it}\right)\right],$$

where the expectation is taken with respect to the process of $x_{it}, \pi_{it}, z_{it}, \nu_{it}$.

Let $V_i(x, z, \pi, \nu)$ denote the value function associated with any given state (x, z, π, ν). Bellman's principle then implies[10]

$$\begin{aligned} V_i(x_t, z_t, \pi_t, \nu_t) = \max_{y_t} \Big\{ & u_i(y_t, x_t, z_t, \nu_t) \\ & + \beta_i \iiint V_i(x_{t+1}, \gamma_i(z_t, x_t, y_t), \pi_{t+1}, \nu_{t+1}) \pi_t(x_{t+1}) \rho_i(\pi_{t+1}; x_{t+1}, \pi_t, x_t) \tau_i(\nu_{t+1}) dx_{t+1} d\pi_{t+1} d\nu_{t+1} \Big\}. \end{aligned} \tag{6}$$

The implied policy rule for actions is then, under suitable regularity conditions (e.g., Stokey, Lucas, and Prescott, 1989),

$$y_{it} = \phi\left(x_{it}, z_{it}, \pi_{it}, \nu_{it}, \rho_i, u_i, \beta_i, \gamma_i, \tau_i\right), \tag{7}$$

for some function ϕ. Let

$$\phi_i\left(x_{it}, z_{it}, \pi_{it}\right) = \int \phi\left(x_{it}, z_{it}, \pi_{it}, \nu_{it}, \rho_i, u_i, \beta_i, \gamma_i, \tau_i\right) \tau_i(\nu_{it}) d\nu_{it}$$

denote the average decision rule with respect to the shocks ν_{it}. It follows from Assumption 1 part 5 that

$$\phi_i\left(x_{it}, z_{it}, \pi_{it}\right) = \mathbb{E}\left[\phi\left(x_{it}, z_{it}, \pi_{it}, \nu_{it}, \rho_i, u_i, \beta_i, \gamma_i, \tau_i\right) \mid x_{it}, z_{it}, \pi_{it}\right].$$

[9] Here π_{it} are not payoff-relevant. However, the nonparametric decision rule below will remain the same if payoffs $u_i(y_{it}, x_{it}, z_{it}, \pi_{it}, \nu_{it})$ depend on π_{it}.

[10] Here the integrals in x_{t+1}, π_{t+1}, and ν_{t+1} are taken relative to appropriate measures.

Hence, (2) holds for $\varepsilon_{it} = y_{it} - \phi_i(x_{it}, z_{it}, \pi_{it})$, which has zero mean given x_{it}, z_{it}, π_{it}. In this framework, ϕ_i in (2) can thus be interpreted as the individual decision rule averaged over the shocks ν_{it}.[11]

Remark 1. *(Finite horizon)*
In a finite horizon environment where $t \in \{1, ..., T_i\}$, the Bellman equation (6) becomes, for $t < T_i$ and some terminal value V_{i,T_i},

$$V_{it}(x_t, z_t, \pi_t, \nu_t) = \max_{y_t} \Big\{ u_i(y_t, x_t, z_t, \nu_t)$$
$$+\beta_i \iiint V_{i,t+1}(x_{t+1}, \gamma_{it}(z_t, x_t, y_t), \pi_{t+1}, \nu_{t+1}) \pi_t(x_{t+1}) \rho_{i,t+1}(\pi_{t+1}; x_{t+1}, \pi_t, x_t) \tau_i(\nu_{t+1}) dx_{t+1} d\pi_{t+1} d\nu_{t+1} \Big\}. \tag{8}$$

Here, differently from Assumption 1, the transitions ρ_{it} between $\pi_{i,t-1}$ and π_{it} are time-specific, and $z_{i,t+1} = \gamma_{it}(z_{it}, x_{it}, y_{it})$. In this case, actions take the form

$$y_{it} = \phi_{it}(x_{it}, z_{it}, \pi_{it}) + \varepsilon_{it},$$

where the dependence of ϕ on i, t stems from the presence of u_i, β_i, τ_i, the terminal value V_{i,T_i}, and the $\rho_{i,s+1}$ and γ_{is} in all periods $s \geqslant t$. An important difference with the infinite-horizon case is that, since ϕ_{it} is time-varying, it is no longer possible to identify individual responses while leaving the individual heterogeneity fully unrestricted.

3.4 Interpreting average partial effects

Structurally interpreting an average partial effect as the effect of a counterfactual change requires two assumptions. The first one is that, as in the framework that we have laid out above, (2) effectively corresponds to the individual's (average) decision rule. The second assumption is that ϕ_i remains invariant in the counterfactual, as we now discuss.

Keeping u_i and β_i constant requires assuming that u_i (such as preferences) and β_i (discounting) are invariant to changes in the environment. This is a common assumption in dynamic structural models. Invariance of the distribution of taste shocks τ_i is also commonly assumed. In turn, keeping γ_i constant requires assuming that the process through which past actions and states feed back onto future z_{it} values is invariant in the counterfactual. When z_{it} is a stock

[11] It is straightforward to include additional state variables s_{it} in (7), under the assumption that beliefs about them are constant and invariant to counterfactual changes. Accounting for additional state variables can be empirically relevant, and we will include a number of such variables as controls in our application.

that depreciates over time or an asset with some return, for example, this requires assuming away the presence of general equilibrium effects through which the return or the depreciation rate might change in the counterfactual.

In addition, as our framework makes clear, structurally interpreting average partial effects requires assuming that the belief updating rule ρ_i remains constant in the counterfactual. A change in ρ_i corresponds to a steady-state or "long-run" counterfactual where the entire process of x_{it}, as perceived by the agent, changes. In our setup, we allow for policies or other counterfactuals to affect beliefs π_{it}, yet we assume that the belief updating rule ρ_i is an individual characteristic that remains unaffected. In Section 7 we describe how to extend the approach to account for beliefs over longer horizons, hence making the invariance assumption about ρ_i less restrictive.

In this paper we study counterfactuals involving changes in x_{it} and π_{it}, while assuming that ρ_i is kept constant. This scenario can be viewed as an intermediate case between a static counterfactual where only x_{it} varies, and a long-run, steady-state counterfactual where the entire long-run belief process, including the belief updating rule ρ_i, is allowed to vary. To identify such long-run counterfactuals in a semi-structural, regression-based approach, one would need to recover the effect of the belief updating rule ρ_i on decisions. This would require the availability of empirical counterparts for ρ_i, as well as suitable cross-sectional exogeneity assumptions (such as the use of an instrument that correlates with ρ_i yet is independent of other structural economic primitives such as preferences). Both conditions would impose strong demands on the data. In particular, ρ_i is a subjective process perceived by the agent, which is not directly informed by responses to subjective expectations questions (since ρ_i need not coincide with the process of realized beliefs π_{it}).

4 Examples

In this section, we describe two examples of our framework. In the first one, we consider a model of consumption, savings, and income, with the aim to assess the effects on consumption of a change in the income process. In the second example, we outline a model of agricultural production that allows farmers to adapt to new climate, with the goal to estimate the effects of current and expected weather.

4.1 Consumption, saving, and income

In the first example, we consider a standard incomplete markets model of consumption and saving behavior. For simplicity, we focus on infinite-horizon environment, as in Chamberlain and Wilson (2000), although the analysis can easily be adapted to a life-cycle environment.

In the model, y_{it} is household i's log consumption in period t, and household utility over consumption is $u_i(y_{it}, \nu_{it})$, where u_i is an increasing utility function and $\nu_{it} \sim \tau_i$ are i.i.d. taste shocks. Household i's discount factor is β_i. Log income x_{it} and beliefs π_{it} about $x_{i,t+1}$ are exogenous and jointly first-order Markov. Households can self-insure using a risk-free bond with constant interest rate r_i, and assets z_{it} follow

$$z_{i,t+1} = (1 + r_i)(z_{it} + w_{it}) - c_{it}, \tag{9}$$

where $w_{it} = \exp(x_{it})$ and $c_{it} = \exp(y_{it})$ denote income and consumption, respectively. As in (7), the (log) consumption rule takes the form[12]

$$y_{it} = \phi \left(x_{it}, \pi_{it}, z_{it}, \nu_{it}, \rho_i, u_i, \beta_i, r_i, \tau_i \right).$$

As a specific example for the income process perceived by the agent, consider a permanent-transitory model (e.g., Hall and Mishkin, 1982):

$$x_{it} = \eta_{it} + u_{it}, \quad \eta_{it} = \eta_{i,t-1} + v_{it},$$

where $u_{it} \sim \mathcal{N}(0, \sigma^2_{iu})$ and $v_{it} \sim \mathcal{N}(0, \sigma^2_{iv})$ are independent over time and independent of each other at all leads and lags. At time t, the agent observes x_{it} and η_{it}, but neither $x_{i,t+1}$ nor $\eta_{i,t+1}$. In this case, we have

$$\pi_{it}(\widetilde{x}) = \frac{1}{\sqrt{\sigma^2_{iu} + \sigma^2_{iv}}} \varphi \left(\frac{\widetilde{x} - \eta_{it}}{\sqrt{\sigma^2_{iu} + \sigma^2_{iv}}} \right), \tag{10}$$

where φ is the standard Gaussian density, and (x_{it}, π_{it}) is first-order Markov. Moreover, the (x_{it}, π_{it}) process induces a stationary conditional distribution $(\pi_{i,t+1} \,|\, x_{i,t+1}, \pi_{it}, x_{it}) \sim \rho_i$. In this specific example, only the mean of π_{it} varies over time and its variance is constant.

Suppose we wish to assess the impact on consumption of a proportional income tax $T(w) = (1 - \delta)w$ introduced at time t, where recall that $w = \exp(x)$ denotes household income. Under the tax, log income is thus $x^{(\delta)} = x + \log \delta$. Suppose households believe the tax change will

[12] Alternatively, in a finite-horizon environment, we obtain a counterpart to this equation, as in Remark 1, which involves a time-varying ϕ_t.

continue being implemented in the future, and they fully adjust their beliefs to the tax. When π_{it} is given by (10) in the absence of the tax, implementing the tax will lead to the new beliefs

$$\pi_{it}^{(\delta)}(\widetilde{x}) = \frac{1}{\sqrt{\sigma_{iu}^2 + \sigma_{iv}^2}} \varphi \left(\frac{\widetilde{x} - \eta_{it} - \log \delta}{\sqrt{\sigma_{iu}^2 + \sigma_{iv}^2}} \right).$$

Hence, the tax affects both the mean of log income and the perceived conditional mean of future log income.

In this model, a proportional tax will not affect the belief updating rule ρ_i.[13] Hence, the total APE fully captures the effect of the tax on consumption. In this case, the contemporaneous APE corresponds to the effect of a purely transitory tax δ at t that will disappear at $t+1$; equivalently, it is the effect of a $\log \delta$-shift in the transitory income shock u_{it}. In turn, the dynamic APE can be interpreted as the effect of a tax that is announced at t and will be implemented at $t+1$.[14] Lastly, the total APE, which is the sum of the contemporaneous and dynamic APEs, corresponds to the effect of a $\log \delta$-shift in the permanent income shock v_{it}.

This model relies on specific assumptions about the income process, information, and beliefs. Those assumptions could be incorrect; for example, agents' expectations may not be rational. In our approach we do not assume that the consumption model with rational expectations and a permanent-transitory income process describes the data. However, a key assumption for interpreting an average partial effect as the structural effect of a counterfactual tax is that, while beliefs π_{it} are affected by the tax, the belief updating rule ρ_i (which matters for long-run income expectations) is not.

Structural and semi-structural tax counterfactuals: a comparison. To illustrate this example, we simulate a large sample from a life-cycle model of consumption and savings based on Kaplan and Violante (2010), where identical, risk-averse households save to smooth consumption while facing borrowing constraints. We study two versions of the model, with rational and adaptive expectations, respectively. Income beliefs, which are key state variables in the model, can be summarized by their time-varying means, which follow a first-order Markov process jointly with log income.

Under both versions of the model, we compute the true effect of a 10% permanent proportional income tax, and we decompose it under the model into a contemporaneous effect due

[13] Indeed, the introduction of the tax is isomorphic to a change in the permanent component, from η_{it} to $\eta_{it}^{(\delta)} = \eta_{it} + \log \delta$. Moreover, the distribution of $(x_{i,t+1}, \eta_{i,t+1})$ given (x_{it}, η_{it}) does not change under the tax.

[14] The DAPE in (4) is evaluated at income $x^{(\delta)}$ after the tax, so that the CAPE and the DAPE add up to the TAPE. It is also possible to compute an alternative DAPE evaluated under income x before the tax, $\widetilde{\Delta}_i^{\text{DAPE}}(\delta, x, \pi, z) = \phi_i(x, \pi^{(\delta)}, z) - \phi_i(x, \pi, z)$.

Table 1: Tax counterfactuals under rational and adaptive expectations

	Rational expectations				Adaptive expectations			
	Structural	Semi-structural			Structural	Semi-structural		
		Linear	Quadratic	Spline		Linear	Quadratic	Spline
CAPE	-0.0163	-0.0151	-0.0150	-0.0150	-0.0122	-0.0344	-0.0191	-0.0133
DAPE	-0.0802	-0.0917	-0.0863	-0.0860	-0.0496	-0.0518	-0.0512	-0.0513
TAPE	-0.0965	-0.1068	-0.1013	-0.1010	-0.0618	-0.0863	-0.0704	-0.0646

Notes: Effects of a 10% permanent income tax on log consumption in two model economies, where households have rational or adaptive expectations, respectively. In both economies, log income follows a permanent-transitory process. For the structural counterfactuals we compute the effect of the tax under the model. For semi-structural ones we regress log consumption on log income, income belief and its interaction with log income, age, age squared, and a function of log assets (linear, quadratic, or 20-knot spline). Households with positive assets, age 26–49.

to current income and a dynamic effect due to beliefs. Then, we compare these counterfactual predictions with our average partial effects (TAPE, CAPE, and DAPE), which we obtain by running consumption regressions in the simulated sample. Since the model has a finite horizon, the consumption function ϕ_t is age-dependent, and we proxy for this dependence by controlling for age and its square. Note that, as we discussed, the belief updating rule ρ_i is invariant under the counterfactual in the rational expectations version of the model. In the adaptive expectations version we assume that invariance is satisfied as well. We provide details about the model, parameter values, and calculation of counterfactuals in Appendix B.

We report the counterfactual calculations in Table 1. Focusing first on the version with rational expectations (in the left panel), the model predicts a decrease in consumption of -0.097, which is almost one-for-one with the tax increase, as is expected in this model, and a large part can be attributed to a change in beliefs. The semi-structural predictions, which do not rely on the knowledge of the structure and parameter values of the structural model but are computed using regressions, come close to these numbers. We report the results of three specifications, where we control for linear, quadratic, or spline functions of log assets, and all of them give comparable results in this case.

Turning next to the version with adaptive expectations (in the right panel), the model predicts a smaller effect of the tax (-0.062), given the expectations process that we assume. When using a structural approach to predict counterfactuals, specifying belief formation correctly is key. However, the semi-structural predictions, which do not rely on correct specification of the model including belief formation, again come close to the tax effects, albeit in this case only when the regression specification is flexible enough (i.e., quadratic or spline).

4.2 Weather and agricultural production

In the second example, we consider a model of production with costly investment. Output $q_{i,t+1} = g_i(x_{i,t+1}, k_{i,t+1})$ depends on the weather $x_{i,t+1}$ and on a dynamic input $k_{i,t+1}$ (such as capital). The weather x_{it}, and farmer i's beliefs π_{it} about $x_{i,t+1}$, are exogenous and jointly first-order Markov. The farmer can invest y_{it} in the dynamic input k_{it} at a cost $c_i(y_{it}, \nu_{it})$, for some i.i.d. cost shifters $\nu_{it} \sim \tau_i$. The dynamic input follows the law of motion $k_{i,t+1} = (1 - \delta_i)k_{it} + y_{it}$. The farmer decides on y_{it} after observing today's weather x_{it} and her beliefs π_{it} about tomorrow's weather, but before observing $x_{i,t+1}$. Lastly, the instantaneous profit in period t is $q_{it} - c_i(y_{it}, \nu_{it})$, and the farmer's discount factor is β_i.

The state variables of the decision problem are x_{it}, π_{it}, k_{it}, and ν_{it}, and, under suitable regularity conditions, the optimal investment rule takes the form

$$y_{it} = \phi\left(x_{it}, \pi_{it}, k_{it}, \nu_{it}, \rho_i, \beta_i, c_i, \delta_i, g_i, \tau_i\right), \tag{11}$$

for some function ϕ. Substituting (11) into the output equation, output in period $t+1$ can thus be written as

$$q_{i,t+1} = \widetilde{\phi}\left(x_{i,t+1}, x_{it}, \pi_{it}, k_{it}, \nu_{it}, \rho_i, \beta_i, c_i, \delta_i, g_i, \tau_i\right), \tag{12}$$

for some function $\widetilde{\phi}$. The presence of π_{it} in (11) and (12) reflects that the farmer may adapt to the prospect of harmful weather in the future by investing today, as emphasized in the literature.[15]

The production function in (12) motivates regressing output on current and past weather and on the weather beliefs. Exploiting changes over time in x_{it} and π_{it}, within farmer, is robust to the presence of individual heterogeneity. As an application, one can use average partial effects to assess the impact of a change in the weather process, which affects both weather realizations and weather beliefs. In this case as well, it is important to note that structurally

[15] For example, Burke and Emerick (2016) rely on a long-difference approach to account for farmers' responses to a changing climate. In turn, Shrader (2020) proposes a framework to account for adaptation in a model where, in contrast with our dynamic framework, the firm's current choice does not affect outcomes (i.e., profit) in later periods. See also Dell, Jones, and Olken (2014) and Keane and Neal (2020). Other approaches rely on specific aspects of the production model, such as envelope condition arguments (Hsiang, 2016, Lemoine, 2018, Gammans, Mérel, Paroissien, et al., 2020).

interpreting the total APE as reflecting the total effect of such a change relies on the assumption that ρ_i, the belief updating process, is invariant. While this assumption may be tenable in the medium run, the total APE will not capture the full impact of long-run changes in the climate under which ρ_i could be affected.

5 Estimating average partial effects

In this section we study identification and estimation of ϕ_i and average partial effects based on model (2).

5.1 Specification and identification

We assume that

$$\mathbb{E}[\varepsilon_{it} \,|\, x_{it}, z_{it}, \pi_{it}] = 0. \tag{13}$$

Note that (13) is satisfied in the structural framework of Section 3. To enhance its plausibility in applications, one can control for additional time-varying regressors (which can be interpreted as additional state variables), as well as for time-invariant fixed-effects. We will account for both factors in our empirical application.[16]

Our approach to the measurement of beliefs π_{it} relies on data about respondents' expectations. Eliciting such responses is becoming increasingly common following the work of Dominitz and Manski (1997), see Manski (2004) for a review. Responses to subjective expectations questions provide information about some features of π_{it}. Typically, the responses can be interpreted as some functionals $m_{it} = m(\pi_{it})$, such as the mean, variance, or some other moments of π_{it}.

We assume that π_{it} is parametrically specified. That is, there exists a finite-dimensional vector θ_{it} such that

$$\pi_{it} = \pi(\cdot; \theta_{it}), \tag{14}$$

where $\pi(\cdot;\theta)$ is known given θ. Parametric specifications are commonly used in the literature on subjective expectations, due to the fact that expectations data are often coarse. For example, in the 1995 and 1998 waves of the SHIW in Italy, respondents are asked about the minimum and maximum earnings that they expect to receive if employed in the following year, together

[16] In certain applications, (13) may not be plausible but one may have access to instruments w_{it} (e.g., that exploit some policy variation in sample) such that $\mathbb{E}[\varepsilon_{it} \,|\, w_{it}] = 0$. Identification of ϕ_i will then hold under suitable relevance conditions (see Newey and Powell, 2003).

with the probability that their earnings will be below the mid-point between those two values. Using these data, Kaufmann and Pistaferri (2009) assume that income beliefs follow a triangular distribution conditional on employment, so that θ_{it} in (14) contains three elements. At the end of this section we will outline how to relax the parametric specification on π_{it}.

Given (2), (13), and (14), we have

$$\phi_i(x_{it}, \pi_{it}, z_{it}) = \mathbb{E}\left[y_{it} \,|\, x_{it}, \theta_{it}, z_{it}\right], \tag{15}$$

so $\phi_i(x, \pi, z)$ is identified for all x, π, z in the empirical support of x_{it}, $\pi_{it} = \pi(\cdot; \theta_{it})$, and z_{it}. In turn, given ϕ_i, average partial effects (TAPE, CAPE and DAPE) are all identified, provided the support of covariates after the change in x_{it} and π_{it} lies within the support of covariates before the change.

The definition of an average partial effect depends on a change in x_{it} and an associated change in beliefs π_{it}. We assume that beliefs remain in the same parametric family after the change, so

$$\pi^{(\delta)} = \pi\left(\cdot; \theta^{(\delta)}\right),$$

for some parameter $\theta^{(\delta)}$. As a benchmark, we assume that the individual fully incorporates the effect of the change from x to $x^{(\delta)}$ in her beliefs, and set

$$\theta^{(\delta)} = \underset{\widetilde{\theta}}{\operatorname{argmax}} \int \log\left(\pi\left(x_{t+1}^{(\delta)}; \widetilde{\theta}\right)\right) \pi(x_{t+1}; \theta) dx_{t+1}. \tag{16}$$

As an example, suppose x is income before a counterfactual tax and $x^{(\delta)} = x + \delta$ is post-tax income. Suppose π_{it} is normal with mean μ_{it} and variance σ_{it}^2, so $\theta_{it} = (\mu_{it}, \sigma_{it}^2)$. Under (16), $\pi_{it}^{(\delta)}$ remains normal after the tax, with mean and variance $\theta_{it}^{(\delta)} = (\mu_{it} + \delta, \sigma_{it}^2)$.

Remark 2. *(Identification in short panels)*
Here we focus on identifying ϕ_i and average partial effects for every individual i, which is relevant for applications with a large time dimension. In short panels, it is not possible to allow for unrestricted individual heterogeneity in ϕ_i. A possible approach is to replace (2) by

$$y_{it} = \phi\left(x_{it}, z_{it}, \pi_{it}\right) + \alpha_i + \varepsilon_{it}, \tag{17}$$

where ϕ is common across individuals, and α_i is an additive individual fixed effect. Under suitable exogeneity assumptions,[17] identification of ϕ can be based on sequential moment restrictions (e.g., Arellano and Bond, 1991). We will follow such an approach in our application.

[17] For example, if (x_{it}, π_{it}) are strictly exogenous and z_{it} are predetermined, one can replace (13) by

$$\mathbb{E}[\varepsilon_{it} \,|\, x_{it}, \pi_{it}, ..., x_{i,1}, \pi_{i,1}, z_{it}, z_{i,t-1}, z_{i,1}] = 0. \tag{18}$$

Remark 3. *(Partial adjustment of beliefs)*
One can define average partial effects associated with other changes in beliefs. For example, assuming that individuals face a cost of adjusting their beliefs that is proportional to the Kullback-Leibler divergence between the beliefs before and after the change, one can replace (16) by

$$\theta^{(\delta)} = \underset{\widetilde{\theta}}{\operatorname{argmax}} \int \log\left(\pi\left(x_{t+1}^{(\delta)}; \widetilde{\theta}\right)\right) \pi(x_{t+1};\theta) dx_{t+1} - \xi \int \log\left(\frac{\pi(x_{t+1};\theta)}{\pi\left(x_{t+1};\widetilde{\theta}\right)}\right) \pi(x_{t+1};\theta) dx_{t+1}. \tag{19}$$

According to (19), $\theta^{(\delta)}$ is given by (16) when the adjustment cost ξ is zero, $\theta^{(\delta)} = \theta$ is unchanged when the cost is infinite, and the individual partially adjust her beliefs for intermediate values of ξ. When available, one could rely on empirical variation in policies and beliefs to discipline ξ. In our application we will focus on the benchmark case $\xi = 0$ where individuals fully adjust their beliefs, while also commenting on the case $\xi = \infty$, which corresponds to the contemporaneous APE where beliefs are held fixed.

5.2 Estimation

For estimation we proceed in three steps. First, we estimate the parameters θ_{it} that govern the belief distribution. Assuming that subjective expectations responses $m_{it} = m(\pi_{it})$ are available, a minimum-distance estimator solves

$$\widehat{\theta}_{it} = \underset{\theta}{\operatorname{argmin}} \ d\left(m_{it}, m(\pi(\cdot;\theta))\right),$$

where d is some distance function (e.g., Euclidean).

In the second step, we estimate ϕ_i as the conditional expectation function in (15). Many approaches are available. For example, Stock (1989) proposes a partially linear semiparametric approach. We will rely on an linear approximation of ϕ_i in a basis of functions,

$$\phi_i(x, \theta, z; \alpha) = \sum_{r=1}^{R} \alpha_{ir} P_r(x, \theta, z), \tag{20}$$

where P_r is a family of functions, such as polynomials, and R is the number of terms. In short panels, we restrict α_{ir} not to depend on i, except the coefficient that corresponds to the intercept in the regression (see Remark 2).

Given observations y_{it}, x_{it}, z_{it} and estimates $\widehat{\theta}_{it}$, for $i = 1, ..., n$ and $t = 1, ..., T$, we estimate α_{ir} using penalized least squares regression

$$\widehat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \sum_{i=1}^{n}\sum_{t=1}^{T} \left(y_{it} - \sum_{r=1}^{R} \alpha_{ir} P_r(x_{it}, \widehat{\theta}_{it}, z_{it})\right)^2 + \operatorname{Pen}(\alpha). \tag{21}$$

In the application we will rely on two choices for the penalty term: no penalty (i.e., $\text{Pen}(\alpha) = 0$) so the estimator is simply OLS, and an ℓ^1 penalty (i.e., $\text{Pen}(\alpha) = \lambda \sum_{i,r} |\alpha_{ir}|$), corresponding to the Lasso estimator.

Lastly, in the third step we estimate counterfactuals by plugging in the estimates $\widehat{\theta}_{it}$ and $\widehat{\alpha}$ in the APE formulas, averaged over individuals and time periods. For example, we estimate the total APE averaged over individuals and time periods as

$$\widehat{\Delta}^{\text{TAPE}}(\delta) = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{r=1}^{R} \widehat{\alpha}_{ir} \left(P_r \left(x_{it}^{(\delta)}, \widehat{\theta}_{it}^{(\delta)}, z_{it} \right) - P_r \left(x_{it}, \widehat{\theta}_{it}, z_{it} \right) \right), \tag{22}$$

with analogous expressions for the contemporaneous and dynamic APEs. When including a large number R of terms in the expansion, and relying on a penalty for regularization, plug-in estimators such as (22) may be biased. To address this issue, in our application we implement the double Lasso method of Belloni, Chernozhukov, and Hansen (2014). Confidence intervals follow readily from standard methods for multi-step estimators.

Remark 4. *(Nonparametric belief distributions)*
The parametric approach we adopt in our application is motivated by the coarse belief information available in the SHIW. In other applications with richer information, a nonparametric treatment of the belief distribution π_{it} may be feasible. Póczos, Singh, Rinaldo, and Wasserman (2013) propose a nonparametric regression estimator that, given a nonparametric estimate $\widehat{\pi}_{it}$, can be used to consistently estimate ϕ_i and average partial effects. However, their estimator suffers from a slow convergence rate in general. An alternative is to assume that ϕ_i in (2) is linear, or more generally polynomial, in beliefs, as in the literature on functional regression (see, e.g., Ramsay and Dalzell, 1991, and Yao and Müller, 2010). Under linearity in beliefs, there exists a function φ_i such that

$$\phi_i(x, z, \pi) = \int \varphi_i(x, z, \widetilde{x}) \pi(\widetilde{x}) d\widetilde{x}, \tag{23}$$

and one can estimate φ_i using functional regression estimators based on principal components analysis or Tikhonov regularization (Hall and Horowitz, 2007). However, all these methods require large samples and the availability of rich information about π_{it}.

Remark 5. *(Bounds)*
Absent parametric assumptions, the information in the subjective expectations responses m_{it} may not be sufficient to point-identify π_{it} nonparametrically. Instead of imposing parametric assumptions, an alternative approach is to follow a partial identification strategy. To illustrate

this approach, let us omit the reference to x and z for conciseness. In this case, the conditional mean $\phi_i(\pi_{it}) = \mathbb{E}[y_{it} \mid \pi_{it}]$ is bounded as follows:

$$\underbrace{\inf_{\pi \in \Pi(m_{it})} \phi_i(\pi)}_{=B_i^L(m_{it};\phi_i)} \leqslant \mathbb{E}[y_{it} \mid \pi_{it}] \leqslant \underbrace{\sup_{\pi \in \Pi(m_{it})} \phi_i(\pi)}_{=B_i^U(m_{it};\phi_i)},$$

where $\Pi(m_{it}) = \{\pi : m(\pi) = m_{it}\}$. These bounds imply the following moment inequalities on ϕ_i:

$$\mathbb{E}\left[y_{it} - B_i^L(m_{it};\phi_i) \mid m_{it}\right] \geqslant 0, \quad \mathbb{E}\left[y_{it} - B_i^U(m_{it};\phi_i) \mid m_{it}\right] \leqslant 0.$$

Remark 6. *(Without expectations data)*
When subjective expectations data are not available, our approach is still applicable provided one can recover estimates of the belief distribution π_{it}. A strategy to do so is to assume that agents have rational expectations, and to make assumptions about the dynamic process of x_{it}. We will return to this point in the extensions section of the paper.

6 Income, consumption, and income expectations

In this section we apply our approach to empirically study how consumption depends on current and expected income, and to conduct various tax counterfactuals.

6.1 Data

The Italian Survey on Household Income and Wealth (SHIW) is a cross-sectional survey that collects information on annual consumption, disposable income, and wealth of Italian families. Since 1989, it includes a panel component. We use the 1989–1991 waves and the 1995–1998 waves, which include questions about income expectations asked to a subsample of households.

The expectations questions differ in both set of waves. However, as we show in Appendix C, the results are qualitatively similar when analyzing the waves separately, so we pool them together to increase power. In 1989 and 1991, individuals are asked about the probability their income growth will fall within a set of predetermined intervals. In 1995 and 1998, individuals are asked the maximum and minimum amounts they expect to earn if employed, and the probability of earning less than the mid-point between the maximum and minimum. We assume beliefs about log income next year follow a normal distribution. In Appendix C we describe our approach to estimate the mean μ_{it} and standard deviation σ_{it} of the beliefs for each individual and time period, which follows Arellano, Bonhomme, De Vera, Hospido, and Wei (2022). We will also comment on robustness checks obtained under different assumptions and estimation strategies.

We focus on employed household heads, while excluding the self-employed. Our cross-sectional sample with information on beliefs has 7,796 household-year observations, and our panel sample with data on beliefs in two consecutive waves for the same head has 1,646 observations. In Appendix Tables C1 and C2 we report descriptive statistics about income expectations questions. In Appendix Table C3 we provide descriptive statistics about income, consumption, assets, and the estimated means and variances of log income beliefs. Belief questions are about individual income, while consumption, assets, and current income are reported at the household level. We will account for this discrepancy in our construction of average partial effects, and we will also report estimates that control for spousal beliefs (when available). Lastly, in Appendix Table C4 we document that beliefs have explanatory power for future log income, even conditional on current log income and other controls.

6.2 Estimates of the consumption function

We estimate several versions of the following regression of log consumption:

$$\begin{aligned} y_{it} =& \phi_i(x_{it}, z_{it}, \pi_{it}) + \varepsilon_{it} \\ =& \beta_x x_{it} + \beta_\pi' \theta_{it} + \beta_{\pi x}' \theta_{it} x_{it} + \beta_z' z_{it} + \alpha_i + \varepsilon_{it}, \end{aligned} \tag{24}$$

where y_{it} is log consumption, x_{it} is log income, θ_{it} contains the mean and variance of income beliefs, and z_{it} include log assets as well as a variety of controls (including age, household composition, and a wave indicator).

We show the estimates in Table 2 where we estimate equation (24) by OLS in first differences. In the table we show standard errors clustered at the household level that do not account for the estimation of the means and variances of beliefs (we will return to the impact of belief elicitation at the end of this subsection). The results in columns (2) and (3) show that the mean income beliefs influence consumption decisions significantly over and beyond current income, while the variance of the beliefs has an insignificant effect. It is interesting to compare the estimates in column (2) with those in column (1) that do not account for beliefs. When including beliefs, the coefficient of family income decreases from 0.58 to 0.44, consistently with the presence of an omitted variable bias in column (1).

In column (4) of Table 2, we interact the mean income beliefs with current income. While the estimates suggest the effect of the mean belief tends to be larger for higher-income households, the interaction effect is only marginally significant and we will see it is not always robust.

Lastly, in column (5) we add the variance of beliefs and its interaction with income, and find small differences compared to column (4), with insignificant coefficients associated with the variance of beliefs. In addition to these specifications we also estimated flexible models using the Lasso, which we will use to produce average partial effects in the next subsection.

In the appendix we report a series of robustness checks. In Appendix Table D1 we probe the robustness of our estimates to different assumptions about the distribution of beliefs, and to different construction methods for the mean and variance of beliefs. In the same table we also show that our estimates are robust to controlling for spouses' beliefs about their own income. In Appendix Table D2 we report estimates for both sets of waves separately. Lastly, in Appendix Table D3 we study the robustness of our estimates to our treatment of assets. Since assets are predetermined (see footnote 17), they are endogenous in first difference, and we use an IV strategy that relies on first-period assets and income as instruments in first difference. A more general issue with assets measurement in the SHIW is that respondents are asked about end-of-year assets, while the argument in the consumption function is beginning-of-period assets. In Appendix Table D3 we also report robustness checks excluding assets in our panel specifications. All the results for current income and income beliefs that we report in the appendix are overall quite similar to those based on our main specification.

Measurement error. A possible concern with the estimates in Table 2 is measurement error in belief data. To explore this issue, we focus on the 1989–1991 waves where individuals are asked to distribute 100 balls into 12 bins, corresponding to different intervals of log income growth (which we assume is normally distributed). A simple model of the responses is that individuals draw 100 i.i.d. values from their normal belief distribution, and put those in the bins. However, we document that, if they were indeed drawing 100 values, respondents would be reporting many more bins on average than they do in the data. As an alternative model, we assume individuals only compute $M < 100$ draws. This could correspond to their inability to imagine a very large number of income growth "scenarios". We show that, when M is lower than 10, the predicted number of bins reported by the individuals is much closer to the data. Given this model of measurement error, for any given M, we implement a "small-σ" approximation (e.g., Evdokimov and Zeleneev, 2022) and use it to bias-correct our regression estimates. We find that, while different M's can imply very different belief responses, the resulting coefficient estimates vary little. We provide details in Appendix E. While this evidence is reassuring, it relies on a specific model of measurement error, and our ability to entertain other models is limited given the short panel dimension available in the SHIW.

Table 2: Estimates of the log consumption function

	(1)	(2)	(3)	(4)	(5)
Mean expected log income		0.235	0.238	0.229	0.231
		(0.094)	(0.095)	(0.093)	(0.093)
(Mean expect. log income)·(Log family income)				0.104	0.104
				(0.061)	(0.061)
Var expected log income			-2.590		-2.613
			(1.876)		(1.941)
(Var expect. log income)·(Log family income)					-1.144
					(3.499)
Log family income	0.584	0.439	0.439	0.439	0.440
	(0.070)	(0.089)	(0.089)	(0.089)	(0.089)
Log family assets	0.010	0.018	0.018	0.019	0.018
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Household fixed effect	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
N observations	1,536	1,536	1,536	1,536	1,536
N households	768	768	768	768	768
R-squared	0.24	0.26	0.26	0.26	0.26
Pvalue F beliefs		0.01	0.03	0.02	0.05

Notes: SHIW, 1989–1991 and 1995–1998. Regression for household heads. The expectations variables (mean and variance) and log family income are centered around the weighted average in the sample. Controls include age and age squared, existence of a spouse, marital status, family size, number of children 0-5, 6-13, 14-17 years old in the household, number of children outside the household, number of income earners in the household, and a wave indicator. Regression results are weighted using survey weights. Standard errors (shown in parenthesis) are clustered at the household level.

6.3 Counterfactual taxes

We now use our framework, and our estimates of the consumption function, to assess the effects of a counterfactual income tax on consumption. We assume that the tax schedule takes the parametric form $T(w_g) = w_g - \lambda w_g^{1-\tau}$, where w_g denotes gross income (e.g., Benabou, 2002). To define a baseline level of the tax, we rely on the estimates obtained by Holter, Krueger, and Stepanchuk (2019) for Italy, averaged over family composition characteristics in our sample.

We consider three counterfactuals corresponding to changes in the λ and τ parameters that index the tax schedule. In the *transitory tax* and *permanent tax* counterfactuals, we increase the average tax by 10 percentage points by decreasing λ, only for one period in the former case and in all subsequent periods in the latter. In the *regressivity* counterfactual, we set the parameter τ to its value in the French tax system, which is less progressive than in Italy, while at the same time decreasing λ such that the tax change is neutral in terms of total tax revenue.

To estimate the effects of the counterfactuals we compute average partial effects. We report estimates of TAPE, CAPE, and DAPE obtained using linear regression (see Table 2), as well as estimates obtained using the Lasso. For the latter, we rely on the double/debiased Lasso method introduced by Belloni, Chernozhukov, and Hansen (2014), based on interactions and powers of the covariates up to the third order. In the calculations for the permanent tax and regressivity counterfactuals, we assume that individuals fully adjust their beliefs to the new tax; i.e., we implement the formula in (16).

The top and bottom panels in Figure 1 show average partial effects based on the estimates from column (5) in Table 2 and on the Lasso, respectively. On the left graphs we show the effects on log consumption of a 10% transitory tax. The overall effect based on OLS is -0.049, and it is very similar according to the Lasso. There is only moderate variation along income quantiles (indicated on the x axis), for both specifications.

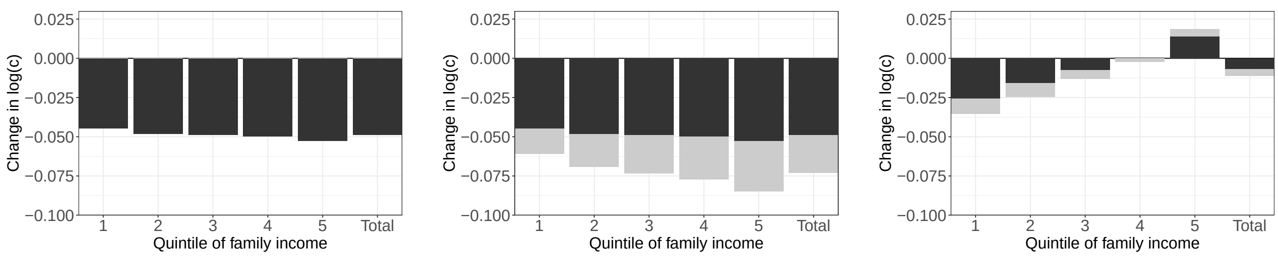
On the middle graphs we show the effect of a 10% permanent tax. Note that the contemporaneous average partial effect (CAPE) corresponds to the effect of a transitory tax (compare with the left graphs). Beyond this contemporaneous effect, dynamic effects are sizable. The dynamic APE (DAPE), which reflects the impact of a changes in beliefs, contributes an additional -0.024 according to OLS, and -0.028 according to the Lasso. The total change in consumption, which is approximately -0.073 in both specifications, is less than the 10% decrease in income, as is expected if households are only partially insured against income changes (Blundell, Pistaferri, and Preston, 2008). Moreover, the estimates from both specifications indicate that dynamic effects are larger for higher-income households.

Lastly, on the right graphs we show the effect of a revenue-neutral decrease in the pro-

Figure 1: Average partial effects for various tax counterfactuals

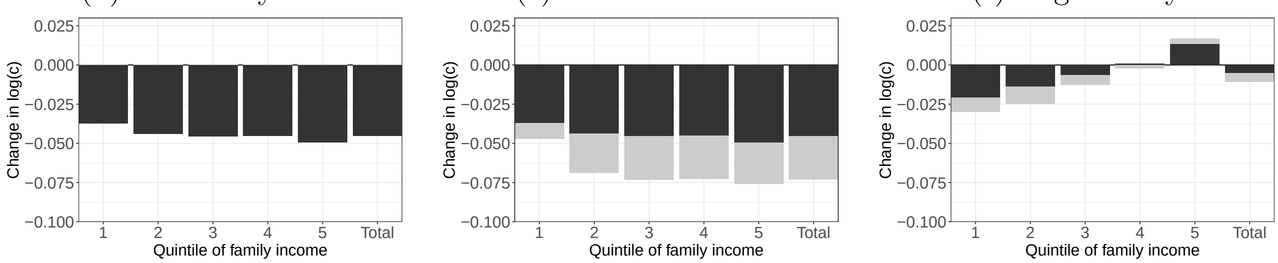
A. Average partial effects based on OLS estimates

(a) Transitory tax (b) Permanent tax (c) Regressivity



B. Average partial effects based on the Lasso

(d) Transitory tax (e) Permanent tax (f) Regressivity



Notes: SHIW, 1989–1991 and 1995–1998, cross-sectional sample. Black bars correspond to contemporaneous APE and grey bars correspond to dynamic APE. Total APE are the sums of CAPE and DAPE. In the top panel we report results based on OLS estimates, see column (5) in Table 2. In the bottom panel we report estimates based on the double/debiased Lasso, for a dictionary including interactions and powers of the covariates up to the third order.

gressivity of the tax. While the total effects averaged over all households are relatively small (around -0.011), they show substantial heterogeneity along the income distribution: reducing progressivity tends to favor the rich, and it hurts the consumption of the poor proportionally more. The estimates of OLS and the Lasso are very similar. Moreover, we observe that, as in the other two counterfactuals, the contemporaneous and dynamic effects of taxes tend to go in the same direction.

It is interesting to compare these estimates with calculations of average partial effects that do not account for the role of beliefs. In that case, the average consumption effect over all households of a 10% permanent income tax is -0.065. This is larger than the contemporaneous effect (-0.049), consistently with beliefs being an omitted relevant regressor in the specification without beliefs. However, this is lower than the total effect that accounts for both contemporaneous and dynamic margins (-0.073). These differences underscore the need to account for beliefs when computing tax counterfactuals. In addition, note that an estimation method that does not include beliefs cannot account for the different impacts between a permanent tax and transitory one.

The estimates in Figure 1 correspond to average partial effects, and our framework provides conditions under which those can truly be interpreted as structural tax counterfactuals. Leaving aside issues about the specification and measurement of beliefs (which we have studied in robustness checks after Table 2), there are two key conditions. The first one is that individual beliefs indeed respond one-to-one to the tax. By varying the parameter ξ in (19) we can predict tax effects under different assumptions about belief responses, in the spirit of sensitivity analysis. The second, crucial condition is that the belief updating rule ρ_i is invariant under the tax. Failure of this assumption does not affect the prediction of the transitory tax increase. However, when tax changes have a long-lasting effect, changes in ρ_i may occur and induce a third margin of response, beyond contemporaneous and dynamic effects. While this third margin may be small or zero in certain cases (as in the permanent-transitory model with a proportional tax in Subsection 4.1), accounting for it may be important in other cases. The extension to beliefs over longer horizons that we outline in the next section provides a possible way forward.

7 Extensions

In this paper we provide a method to account for the role of individual expectations in assessing the impact of policies and other counterfactuals. Our approach is semi-structural, in the sense that it is justified under dynamic structural assumptions, yet implementing the method does not require full specification and estimation of a structural model.

Among possible extensions of the method, it is interesting to allow for state-contingent beliefs. For example, in a model of occupational choice, individual income beliefs contingent on occupational choice may be available (e.g., Patnaik, Venator, Wiswall, and Zafar, 2020, Arcidiacono, Hotz, Maurel, and Romano, 2020). In that case, the state-contingent beliefs enter as arguments in the decision rule. A second extension is to introduce beliefs over longer horizons. As an extreme case, if one had access to data on the sequence of beliefs about $x_{i,t+1}, x_{i,t+2}, ...$ into the far future, accounting for those as determinants of the decision, and shifting them in the counterfactual, would provide valid predictions without the need for an invariance assumption about some ρ_i process. To go one step in this direction, one can elicit beliefs over multiple horizons $x_{i,t+1}, x_{i,t+2}, ..., x_{i,t+S}$ (Koşar and van der Klaauw, 2022), and account for variation in those beliefs in estimation and counterfactuals. In Appendix G we further discuss these two extensions. Moreover, extending the framework to allow beliefs to be endogenous, in the sense that past actions may shape future beliefs, will be an important task for future work.

Lastly, we have shown how to use data on subjective expectations to estimate individual beliefs. While such data are increasingly available, belief estimates could be obtained in other ways. In ongoing work, we apply the method to weather and corn yields data since 1950. In this case we do not have data on farmers' beliefs, and, to estimate the beliefs, we assume that farmers have rational expectations and that the weather follows a hidden Markov stochastic process. We find that beliefs affect agricultural output, and that dynamic responses to a change in the weather process, which may reflect adaptation, tend to undo part of the contemporaneous effect.

References

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020): "Ex ante returns and occupational choice," *Journal of Political Economy*, 128(12), 4475–4522.

Arellano, M., R. Blundell, and S. Bonhomme (2017): "Earnings and consumption dynamics: a nonlinear panel data framework," *Econometrica*, 85(3), 693–734.

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(2), 277–297.

Arellano, M., S. Bonhomme, M. De Vera, L. Hospido, and S. Wei (2022): "Income risk inequality: Evidence from Spanish administrative records," *Quantitative Economics*, 13(4), 1747–1801.

Armona, L., A. Fuster, and B. Zafar (2019): "Home price expectations and behaviour: Evidence from a randomized information experiment," *The Review of Economic Studies*, 86(4), 1371–1410.

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., A. Kovacs, and K. Molnar (2020): "Euler equations, subjective expectations and income shocks," *Economica*, 87(346), 406–441.

Bachmann, R., G. Topa, and W. van der Klaauw (2022): *Handbook of Economic Expectations*. Elsevier.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 81(2), 608–650.

Benabou, R. (2002): "Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency?," *Econometrica*, 70(2), 481–517.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *American Economic Review*, 98(5), 1887–1921.

Bover, O. (2015): "Measuring expectations from household surveys: new results on subjective probabilities of future house prices," *SERIEs*, 6(4), 361–405.

Burke, M., and K. Emerick (2016): "Adaptation to climate change: Evidence from US agriculture," *American Economic Journal: Economic Policy*, 8(3), 106–40.

Chamberlain, G., and C. A. Wilson (2000): "Optimal intertemporal consumption under uncertainty," *Review of Economic Dynamics*, 3(3), 365–395.

Chetverikov, D., Z. Liao, and V. Chernozhukov (2021): "On cross-validated lasso in high dimensions," *The Annals of Statistics*, 49(3), 1300–1317.

Deaton, A. (1992): *Understanding consumption*. Oxford University Press.

Delavande, A. (2008): "Pill, patch, or shot? Subjective expectations and birth control choice," *International Economic Review*, 49(3), 999–1042.

Delavande, A., X. Giné, and D. McKenzie (2011): "Measuring subjective expectations in developing countries: A critical review and new evidence," *Journal of Development Economics*, 94(2), 151–163.

Dell, M., B. F. Jones, and B. A. Olken (2014): "What do we learn from the weather? The new climate-economy literature," *Journal of Economic Literature*, 52(3), 740–98.

Deschênes, O., and M. Greenstone (2007): "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather," *American Economic Review*, 97(1), 354–385.

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

——— (2007): "Expected equity returns and portfolio choice: Evidence from the Health and Retirement Study," *Journal of the European Economic Association*, 5(2-3), 369–379.

Evdokimov, K. S., and A. Zeleneev (2022): "Simple estimation of semiparametric models with measurement errors," Discussion paper, cemmap working paper.

Gammans, M., P. Mérel, E. Paroissien, et al. (2020): "Reckoning climate change damages along an envelope," in *2020 Annual Meeting, July 26-28, Kansas City, Missouri*, no. 304475. Agricultural and Applied Economics Association.

Gouveia, M., and R. P. Strauss (1994): "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," *National Tax Journal*, 47(2), 317–339.

Guiso, L., T. Jappelli, and L. Pistaferri (2002): "An empirical analysis of earnings and employment risk," *Journal of Business & Economic Statistics*, 20(2), 241–253.

Guiso, L., and G. Parigi (1999): "Investment and demand uncertainty," *The Quarterly Journal of Economics*, 114(1), 185–227.

Hall, P., and J. L. Horowitz (2007): "Methodology and convergence rates for functional linear regression," *The Annals of Statistics*, 35(1), 70–91.

Hall, R. E., and F. S. Mishkin (1982): "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 50(2), 461–481.

Holter, H. A., D. Krueger, and S. Stepanchuk (2019): "How do tax progressivity and household heterogeneity affect Laffer curves?," *Quantitative Economics*, 10(4), 1317–1356.

Hsiang, S. (2016): "Climate econometrics," *Annual Review of Resource Economics*, 8, 43–75.

Hurd, M. D., J. P. Smith, and J. M. Zissimopoulos (2004): "The effects of subjective survival on retirement and social security claiming," *Journal of Applied Econometrics*, 19(6), 761–775.

Ichimura, H., and C. Taber (2002): "Semiparametric reduced-form estimation of tuition subsidies," *American Economic Review*, 92(2), 286–292.

Ichimura, H., and C. R. Taber (2000): "Direct Estimation of Policy Impacts," Working Paper 254, National Bureau of Economic Research.

Kaplan, G., and G. L. Violante (2010): "How Much Consumption Insurance beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2(4), 53–87.

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

Keane, M., and T. Neal (2020): "Climate change and US agriculture: Accounting for multidimensional slope heterogeneity in panel data," *Quantitative Economics*, 11(4), 1391–1429.

Keane, M. P., and K. I. Wolpin (2002a): "Estimating welfare effects consistent with forward-looking behavior. Part I: Lessons from a simulation exercise," *Journal of Human Resources*, pp. 570–599.

——— (2002b): "Estimating Welfare Effects Consistent with Forward-Looking Behavior. Part II: Empirical Results," *Journal of Human Resources*, pp. 600–622.

Koşar, G., and C. O'Dea (2022): "Expectations Data in Structural Microeconomic Models," Working Paper 30094, National Bureau of Economic Research.

Koşar, G., and W. van der Klaauw (2022): "Workers' Perceptions of Earnings Growth and Employment Risk," .

Lemoine, D. (2018): "Estimating the consequences of climate change from variation in weather," Working Paper 25008, National Bureau of Economic Research.

Lochner, L. (2007): "Individual perceptions of the criminal justice system," *American Economic Review*, 97(1), 444–460.

Lucas, R. (1976): "Econometric policy evaluation: a critique, w: The Phillips Curve and Labor Markets, (ed.: K. Brunner and AH Meltzer)," .

Manski, C. F. (2004): "Measuring expectations," *Econometrica*, 72(5), 1329–1376.

Manski, C. F., and F. Molinari (2010): "Rounding probabilistic expectations in surveys," *Journal of Business & Economic Statistics*, 28(2), 219–231.

Marschak, J. (1953): "Economic Measurements for Policy and Prediction. Pp. 1-26 in Studies in Econometric Method, ed. WC Hood and TP Koopmans," .

Newey, W. K., and J. L. Powell (2003): "Instrumental variable estimation of nonparametric models," *Econometrica*, 71(5), 1565–1578.

Patnaik, A., J. Venator, M. Wiswall, and B. Zafar (2020): "The role of heterogeneous risk preferences, discount rates, and earnings expectations in college major choice," *Journal of Econometrics*.

Pesaran, M. H., and M. Weale (2006): "Survey expectations," *Handbook of Economic Forecasting*, 1, 715–776.

PISTAFERRI, L. (2001): "Superior information, income shocks, and the permanent income hypothesis," *Review of Economics and Statistics*, 83(3), 465–476.

PÓCZOS, B., A. SINGH, A. RINALDO, AND L. WASSERMAN (2013): "Distribution-free distribution regression," in *Artificial Intelligence and Statistics*, pp. 507–515. PMLR.

RAMSAY, J. O., AND C. DALZELL (1991): "Some tools for functional data analysis," *Journal of the Royal Statistical Society: Series B (Methodological)*, 53(3), 539–561.

SHRADER, J. (2020): "Improving Climate Damage Estimates by Accounting for Adaptation," *Available at SSRN 3212073*.

STINEBRICKNER, R., AND T. R. STINEBRICKNER (2014): "A major in science? Initial beliefs and final outcomes for college major and dropout," *Review of Economic Studies*, 81(1), 426–472.

STOCK, J. H. (1989): "Nonparametric policy analysis," *Journal of the American Statistical Association*, 84(406), 567–575.

STOKEY, N. L., R. E. LUCAS, AND E. PRESCOTT (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

STOLTENBERG, C. A., AND A. UHLENDORFF (2022): "Consumption Choices and Earnings Expectations: Empirical Evidence and Structural Estimation," .

VAN DER KLAAUW, W. (2012): "On the use of expectations data in estimating structural dynamic choice models," *Journal of Labor Economics*, 30(3), 521–554.

VAN DER KLAAUW, W., AND K. I. WOLPIN (2008): "Social security and the retirement and savings behavior of low-income households," *Journal of Econometrics*, 145(1-2), 21–42.

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

WOLPIN, K. I. (2013): *The limits of inference without theory*. MIT Press.

YAO, F., AND H.-G. MÜLLER (2010): "Functional quadratic regression," *Biometrika*, 97(1), 49–64.

APPENDIX

A A belief formation model with learning

In this section of the appendix we describe a model of belief formation with learning that we mentioned in Subsection 3.2 in the main text. Suppose that

$$x_{it} = \alpha_i + \varepsilon_{it},$$

where ε_{it} are i.i.d. $\mathcal{N}(0, \sigma^2_{\varepsilon_i})$. Suppose agents have rational expectations, with information set $\Omega_{it} = \{x_i^t\}$, which does not include α_i. Furthermore, assume agents are Bayesian learners with prior beliefs about α_i that are normally distributed. Then, by Bayes rule, posterior beliefs about α_i over time are also normally distributed with mean μ_{it} and variance σ^2_{it} satisfying

$$\mu_{it} = \mu_{i,t-1} + \frac{\sigma^2_{it}}{\sigma^2_{\varepsilon_i}}\Big(x_{it} - \mu_{i,t-1}\Big), \tag{A1}$$

$$(\sigma^2_{it})^{-1} = (\sigma^2_{i,t-1})^{-1} + (\sigma^2_{\varepsilon_i})^{-1}. \tag{A2}$$

Then, π_{it} is a normal density with mean $\mathbb{E}_{\pi_{it}}(x_{i,t+1}) = \mu_{it}$ and variance $\mathrm{Var}_{\pi_{it}}(x_{i,t+1}) = \sigma^2_{it} + \sigma^2_{\varepsilon_i}$. Hence, by (A1)-(A2) the belief process satisfies

$$\pi_{it} \,|\, \pi_i^{t-1}, x_i^t \sim \pi_{it} \,|\, \pi_{i,t-1}, x_{it},$$

which is compatible with our framework. Note that the mean beliefs in (A1) are as in the adaptive expectations case, see (5), but with a parameter $\lambda_{it} = \frac{\sigma^2_{it}}{\sigma^2_{\varepsilon_i}}$ that is time-varying and converges to zero over time.

B Structural and semi-structural counterfactuals

In this section of the appendix we present the details of the calibration that we used to produce Table 1, and report additional output from the simulation.

B.1 Model

The model closely follows Kaplan and Violante (2010), with some differences. Agents live for T periods, and work until age T_{ret}, where both T and T_{ret} are exogenous and fixed. *Ex ante* identical households maximize expected life-time utility

$$\mathbb{E}_0\left[\sum_{t=1}^{T} \beta^{t-1} u(c_{it})\right].$$

During working years $1 \leqslant t \leqslant T_{\text{ret}}$, agents receive after-tax labor income w_{it}, the log of which is the sum of a deterministic experience profile κ_t, a permanent component η_{it}, and a transitory component ε_{it},

$$\begin{aligned} x_{it} &= \kappa_t + \underbrace{\eta_{it} + \varepsilon_{it}}_{=x_{it}^{\text{stoch}}}, \\ \eta_{it} &= \eta_{it-1} + v_{it}, \end{aligned}$$

where η_{i1} is drawn from an initial normal distribution with mean zero and variance $\sigma^2_{\eta_1}$. The shocks ε_{it} and v_{it} have zero mean, are independent at all leads and lags, and are normally distributed with variances σ^2_ε and σ^2_v, respectively.

We define gross labor income as $\widetilde{w}_{it} = G(w_{it})$, where G function is the inverse of the tax function[1]

$$\tau(\widetilde{w}_{it}) = \widetilde{w}_{it} - \tilde{\lambda}\widetilde{w}_{it}^{1-\tau}.$$

After retirement, agents receive after-tax social security transfers w_{it}^{ss}, which are a function of average individual gross income over the last few years of their working life,

$$w_{it}^{\text{ss}} = P\left(\frac{1}{T_{\text{ret}} - T_{\text{cont}}} \sum_{t=T_{\text{cont}}}^{T_{\text{ret}}-1} \widetilde{w}_{it}\right).$$

Lastly, throughout their lifetime households can save (but not borrow) through a single risk-free, one-period bond whose constant return is $1 + r$, and they face a period-to-period budget constraint

$$\begin{aligned} z_{i,t+1} &= (1+r)z_{it} + w_{it} - c_{it} \quad \text{if } t < T_{\text{ret}} \\ z_{i,t+1} &= (1+r)z_{it} + w_{it}^{\text{ss}} - c_{it} \quad \text{if } t \geqslant T_{\text{ret}}. \end{aligned}$$

We consider two cases:

- A case with rational expectations, where individuals observe η_{it} each period, and beliefs about after-tax log income next period are normally distributed,

$$\begin{aligned} \mathbb{E}_t(x_{i,t+1}) &= \kappa_{t+1} + \eta_{it}, \\ \text{Var}_t(x_{i,t+1}) &= \sigma^2_v + \sigma^2_\varepsilon. \end{aligned}$$

[1] Kaplan and Violante (2010) use instead a tax function estimated by Gouveia and Strauss (1994).

- A case with adaptive expectations, where beliefs about after-tax log income next period are normally distributed,

$$\begin{aligned}\mathbb{E}_t(x_{i,t+1}) &= \kappa_{t+1} + (\mathbb{E}_{t-1}(x_{it}) - \kappa_t) + \Gamma \cdot (x_{it} - \mathbb{E}_{t-1}(x_{it})) + u_{it}, \quad u_{it} \sim \mathcal{N}(0, V_u), \\ \mathrm{Var}_t(x_{i,t+1}) &= \sigma_v^2 + \sigma_\varepsilon^2,\end{aligned}$$

where Γ is a constant, u_{it} are independent of all other shocks in the model, and initial mean beliefs are given by $\mathbb{E}_1(x_{i2}) = \kappa_2 + \eta_{i1}$.

B.2 Calibration

We closely follow the preferred calibration parameters in Kaplan and Violante (2010) (see also the related calibration in Arellano, Blundell, and Bonhomme, 2017).

Demographics. The model period is one year. Agents enter the labor market at age 25, retire at age 60, and die with certainty at age 95. So we set $T_{\text{ret}} = 35$, and $T = 70$.

Preferences. The utility function is CRRA, $u(c) = c^{1-\gamma}/(1-\gamma)$, where the risk aversion parameter is set to $\gamma = 2$.

Discount factor and interest rate. The interest rate is $r = 0.03$, and $\beta = 1/(1+r)$.

Income process. We use the deterministic age profile κ_t from Arellano, Blundell, and Bonhomme (2017). For the stochastic components of the income process, we set $\sigma_{\eta_1}^2 = 0.15$, $\sigma_v^2 = 0.01$, and $\sigma_\varepsilon^2 = 0.05$.

Initial wealth and borrowing limit. Households' initial assets are set to 0 and there is no borrowing possible.

Tax system. We use parameters derived from Holter, Krueger, and Stepanchuk (2019), $\widetilde{\lambda} = 3.826$, $\tau = 0.137$.

Social security benefits. Social security benefits are a function of average individual gross earnings between the ages of 50 and 60, $w_{it}^{\text{ss}} = P\left(\frac{1}{T_{\text{ret}} - T_{\text{cont}}} \sum_{t=T_{\text{cont}}}^{T_{\text{ret}}-1} \widetilde{w}_{it}\right)$, where $T_{\text{cont}} = 25$. Pre-tax benefits are equal to 90% of average past earnings up to a given bend point, 32% from this first bend point to a second bend point, and 15% beyond that. The two bend points are set at, respectively, 0.18 and 1.10 times cross-sectional average gross earnings. Benefits are then scaled proportionately so that a worker earning average labor between ages 50 and 60 is entitled to a pre-tax replacement rate of 45%. There is also a cap on pre-tax earnings contributing to pensions (cap of 2.2) and only 85% of pre-tax pensions are taxed.

Adaptive beliefs. We take $\Gamma = 0.5$ and $V_u = 0.2$.

There are two main differences between our calibration and the one from Kaplan and Violante (2010), besides including the adaptive expectations case and using a different tax function and income age profile. First, pensions depend on contributions made between ages 50 and 60, so the history of past income is not a relevant state variable before age 50. Second, we do not consider random mortality during retirement years.

B.3 Additional simulation results

In this subsection we report results based on the calibrated structural model that we introduced in Subsection 4.1.

In Table B1 we report structural and semi-structural counterfactual effects of a permanent 10% income tax, as in Table 1, for three different ages: 26, 35, and 45. We see that, under rational expectations (left column), the contemporanous effect of the tax is higher for the young than for older households, while the dynamic impact is lower. This reflects the fact that households start their working life without assets, and that they cannot borrow. The semi-structural average partial effects reproduce the structural policy effects well. In the case of adaptive expectations (right column) there is less variation by age, and while a linear specification tends to produce too high a contemporaneous effect for the old, the quadratic and spline specifications agree well with the structural predictions. For completeness, in Figures B1 and B2 we plot the policy rules and the mean and variance profiles of consumption, assets and income under the model.

Table B1: Simulated tax counterfactuals under rational and adaptive expectations by age

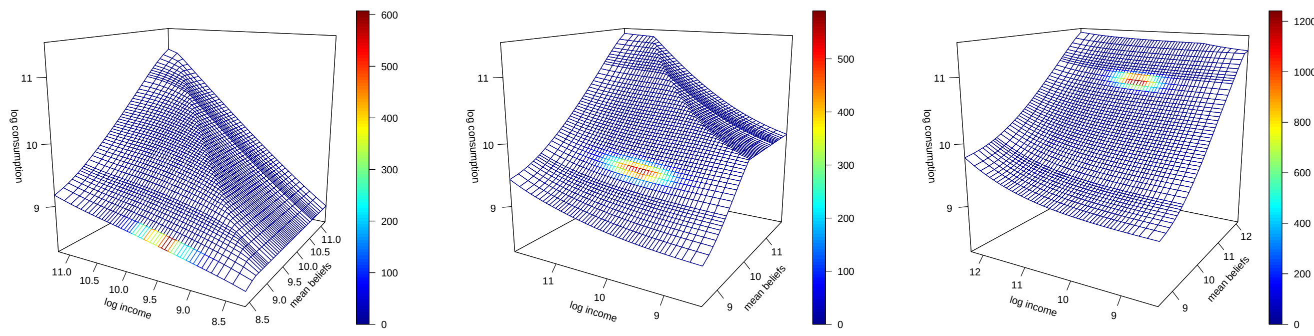
	Rational expectations				Adaptive expectations			
	Structural	Semi-structural			Structural	Semi-structural		
		Linear	Quadratic	Spline		Linear	Quadratic	Spline
Age 26								
CAPE	-0.0663	-0.0599	-0.0599	-0.0599	-0.0331	-0.0318	-0.0313	-0.0315
DAPE	-0.0471	-0.0550	-0.0543	-0.0540	-0.0509	-0.0536	-0.0536	-0.0535
TAPE	-0.1134	-0.1149	-0.1142	-0.1139	-0.0840	-0.0854	-0.0849	-0.0850
Age 35								
CAPE	-0.0110	-0.0097	-0.0097	-0.0097	-0.0111	-0.0284	-0.0149	-0.0123
DAPE	-0.0921	-0.0982	-0.0948	-0.0945	-0.0507	-0.0521	-0.0519	-0.0519
TAPE	-0.1031	-0.1079	-0.1044	-0.1041	-0.0618	-0.0805	-0.0668	-0.0643
Age 45								
CAPE	-0.0058	-0.0062	-0.0062	-0.0061	-0.0078	-0.0337	-0.0139	-0.0084
DAPE	-0.0794	-0.0877	-0.0821	-0.0805	-0.0479	-0.0508	-0.0490	-0.0491
TAPE	-0.0852	-0.0939	-0.0883	-0.0866	-0.0557	-0.0846	-0.0629	-0.0575

Notes: See the notes to Table 1. Results by age.

Figure B1: Policy rules by type of expectations and age

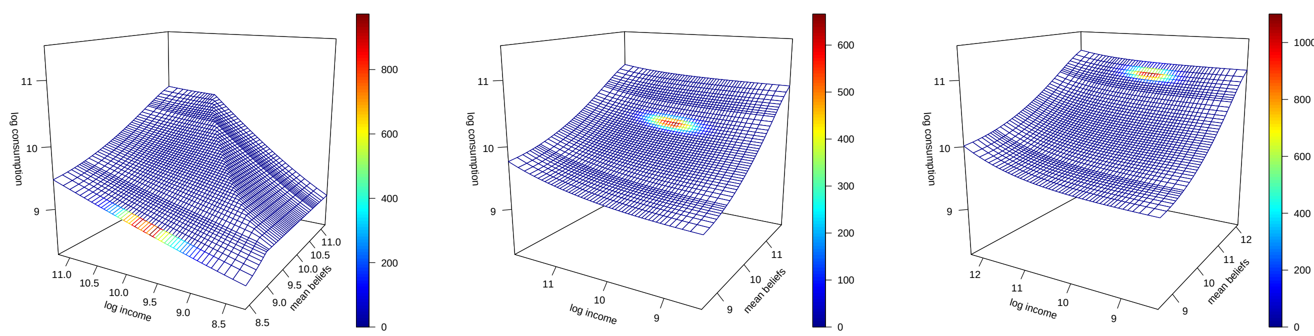
A. Rational expectations

(a) 26 years old (b) 35 years old (c) 45 years old



B. Adaptive expectations

(a) 26 years old (b) 35 years old (c) 45 years old

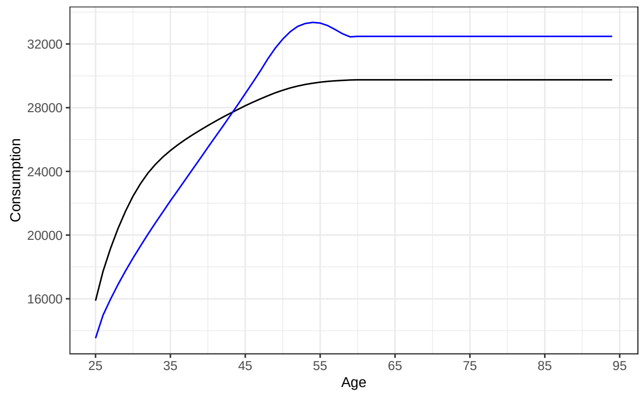


Notes: The top panel plots policy rules under rational expectations and the bottom panel plots policy rules under adaptive expectations. In each figure, assets are fixed at the median value among simulated cases with positive assets. The colors represent the number of observations in the corresponding simulated data set.

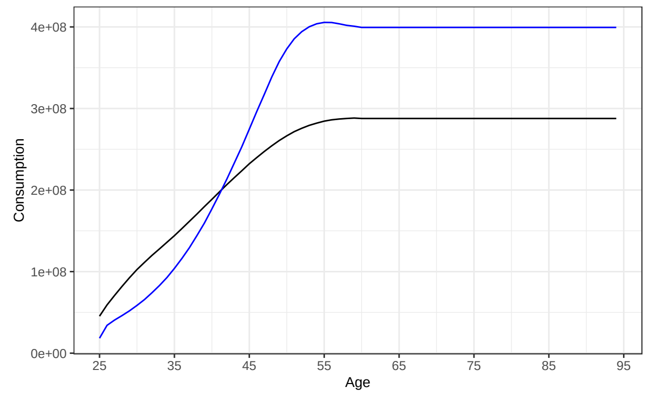
Figure B2: Simulation results, rational versus adaptive expectations

A. Consumption

(a) Mean

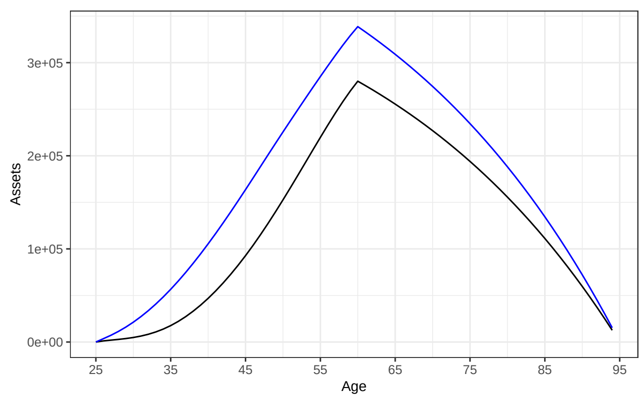


(b) Variance

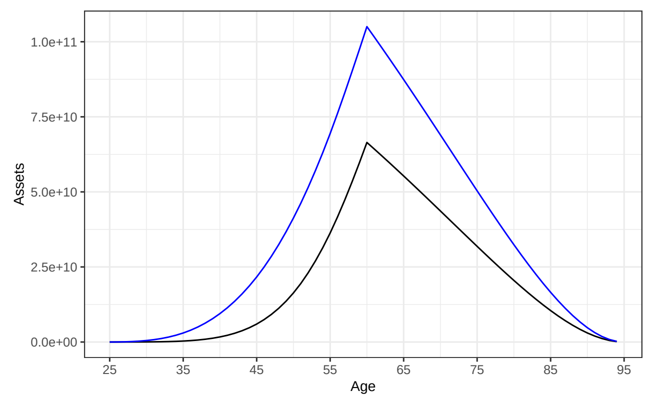


B. Assets

(a) Mean

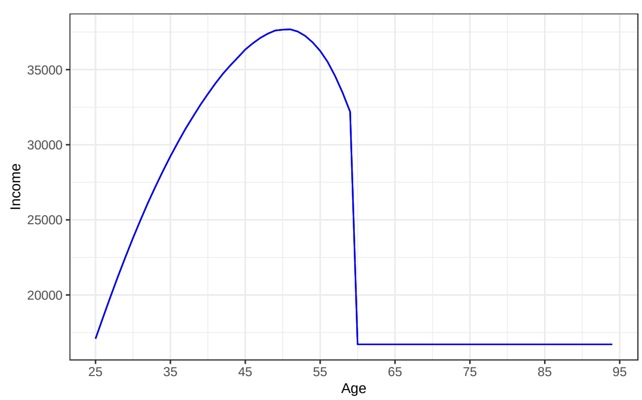


(b) Variance

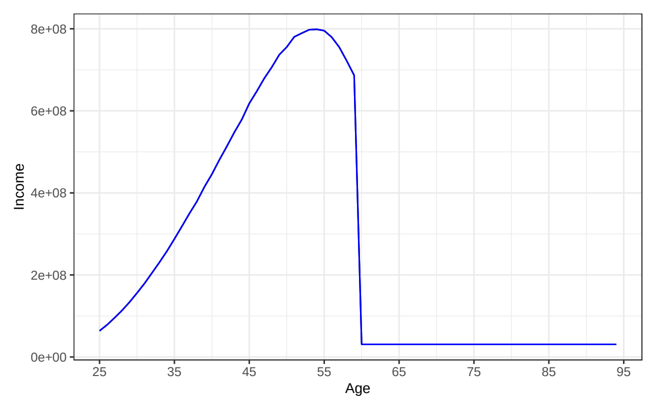


B. Income

(a) Mean



(b) Variance



Notes: Simulations results based on the structural model. Black lines are results under rational expectations, blue lines are results under adaptive expectations.

C Beliefs data

In this section of the appendix we describe the income belief questions in the SHIW, and explain how we estimate the parameters of the belief distributions.

C.1 Expectations questions in the SHIW

The SHIW includes questions about income expectations in waves 1989–1991 and 1995–1998; however the expectations questions differ in the two sets of waves.

The 1989–1991 waves include a question about expected income growth:
Thinking now of your total income from work or retirement and its evolution [for the next 12 months]... Which categories would you exclude? Suppose you have 100 points to distribute among the remaining categories, how many would you give to each?

The possible categories are more than 25%, between 20% and 25%, between 15% and 20%, between 13% and 15%, between 10% and 13%, between 8% and 10%, between 7% and 8%, between 6% and 7%, between 5% and 6%, between 3% and 5%, between 0% and 3%, or less than 0%, and in that case, by how much. In Table C1 we report descriptive statistics corresponding to this question.

The 1995–1998 waves include three questions about expected income level:
Minimum amount expected to earn: *Assuming that you remain in or find employment in the next 12 months, can you say what is the minimum overall ANNUAL amount you expect to earn, net of taxes, including overtime, bonuses, fringe benefits, etc?*
Maximum amount expected to earn: *Assuming again that you remain in or find employment in the next 12 months, can you say what is the maximum overall ANNUAL amount you expect to earn, net of taxes, including overtime, bonuses, fringe benefits, etc?*
Probability of earning less than half: *What is the probability that you will earn less than X (the amount obtained for (MAXIMUM + MINIMUM)/2 ? If you had to give a score of between 0 and 100 to the chances of earning less than X, what would it be? ("0" if certain of earning more than X, "100" if certain of earning less than X).*

In Table C2 we report descriptive statistics corresponding to these questions. In these two waves, the survey also includes a question about the probability of being employed next year that we use in a robustness check specific to those waves.

Table C1: Descriptive statistics on income expectations questions 1989–1991

	Cross-sectional sample				Panel sample			
	Obs	P25	Mean	P75	Obs	P25	Mean	P75
Income growth > 25%	5,486	0	0.79	0	1,096	0	0.63	0
Income growth $20 - 25\%$	5,486	0	0.85	0	1,096	0	1.18	0
Income growth $15 - 20\%$	5,486	0	1.80	0	1,096	0	1.09	0
Income growth $13 - 15\%$	5,486	0	2.72	0	1,096	0	2.92	0
Income growth $10 - 13\%$	5,486	0	5.50	0	1,096	0	4.85	0
Income growth $8 - 10\%$	5,486	0	8.22	0	1,096	0	8.50	0
Income growth $7 - 8\%$	5,486	0	6.78	0	1,096	0	7.99	0
Income growth $6 - 7\%$	5,486	0	7.70	0	1,096	0	9.01	0
Income growth $5 - 6\%$	5,486	0	12.18	0	1,096	0	13.15	5
Income growth $3 - 5\%$	5,486	0	20.49	30	1,096	0	20.16	30
Income growth $0 - 3\%$	5,486	0	29.24	80	1,096	0	28.13	70
Income growth < 0%	5,486	0	3.72	0	1,096	0	2.39	0
Income growth - by how much if < 0%	163	3	10.05	10	15	1	12.18	12

Notes: Descriptive statistics are weighted using the survey's weights.

Table C2: Descriptive statistics on income expectations questions 1995–1998

	Cross-sectional sample				Panel sample			
	Obs	P25	Mean	P75	Obs	P25	Mean	P75
Minimum amount expected to earn	2,310	13,515.1	18,401.7	20,503.5	550	14,645.4	18,866.1	21,968.1
Maximum amount expected to earn	2,310	16,109.9	21,363.3	23,798.7	550	16,893.8	21,551.2	24,897.1
Prob. of earning less than half	2,302	40.00	50.73	70.00	548	30.00	50.75	70.00

Notes: Amounts are in 2010 euros. Descriptive statistics are weighted using the survey's weights.

C.2 Estimation of income beliefs

We assume log income beliefs are normally distributed, with mean μ_{it} and variance σ_{it}^2, and use the expectations questions to estimate these two parameters for each individual and wave. In this subsection, we omit the reference to i and t for ease of notation.

First two waves. For the 1989–1991 waves, we use the survey expectations questions to estimate the mean and variance of the beliefs of log income growth, which are normally distributed under our assumptions, with mean $\mu_g = \mu - x$ (where x is the current log income), and $\sigma_g^2 = \sigma^2$. Given estimates of μ_g and σ_g^2, we then recover estimates of μ and σ^2.

Let $\widehat{p}_j$ denote the fraction of points the respondent assigns to bin j (out of 100 points), for $j = 1, ..., J$, where $J = 12$. For each bin, one could interpret $\widehat{p}_j$ as the probability that a $\mathcal{N}(\mu_g, \sigma_g^2)$ draw takes values within the interval corresponding to that bin. Under this interpretation, one could estimate μ_g and σ_g using maximum likelihood or minimum distance given the frequencies $\widehat{p}_j$. However, this approach does not work well in practice since many of the $\widehat{p}_j$'s are exactly 0 or 1.

Instead of assuming that respondents report exact, normality-based probabilities, we follow Arellano, Bonhomme, De Vera, Hospido, and Wei (2022) and assume that, when answering the survey expectations questions, individuals sample draws from their underlying $\mathcal{N}(\mu_g, \sigma_g^2)$ distribution, and use those draws to provide their answers $\widehat{p}_j$. Given that, in the survey, individuals are asked to distribute 100 points among the 12 bins, we take $M = 100$ as our baseline. Hence, the answers $\widehat{p}_j$ are obtained from $M = 100$ trials from a multinomial distribution with true probabilities p_j.

To estimate the p_j, we assume an uninformative (Jeffreys) prior on $(p_1, ..., p_J)$. It then follows that the posterior means of the p_j are

$$\widetilde{p}_j \quad = \quad \frac{\widehat{p}_j + \frac{1}{2M}}{1 + \frac{J}{2M}}, \quad j = 1, ..., J. \tag{A3}$$

The estimates $\widetilde{p}_j$ are regularized counterparts to the $\widehat{p}_j$. An advantage is that they take values in the open interval $(0, 1)$, so minimum distance or maximum likelihood estimation strategies based on them are well defined. We have performed robustness checks using other regularization devices, including different value for M, and found only minor impacts on the results (see Section D of this appendix).

Given the regularized responses $\widetilde{p}_j$ in (A3), we then construct the cumulative probabilities, $\widetilde{c}_j = \sum_{k=1}^{j} \widetilde{p}_k$, and estimate μ_g and σ_g based on the following system of linear equations:

$$\Phi^{-1}(\widetilde{c}_j) \cdot \sigma_g + \mu_g \quad = \quad v_j, \quad j = 1, ..., J - 1, \tag{A4}$$

where v_j correspond to the right endpoint of the j-th bin, and Φ denotes the standard normal cdf. Since the first and last bins in the survey question are unbounded, we add bounds to those (-10% for the bin below 0%, and 35% for the bin above 25%).[2] This amounts to working with 14 bins in total. We then estimate μ_g and σ_g using OLS based on a subset of the equalities in (A4). Specifically, we use all the bins j for which $\widehat{p}_j > 0$, and use in addition one unbounded bin to the left and one unbounded bin to the right of those. The reason for only using a subset of the restrictions in (A4) is to reduce the influence of the regularization for bins with $\widehat{p}_j = 0$.[3]

As an example, consider an individual who assigns 60 points to the 5–6% bin, and 40 points to the 6–7% bin. In this case we use the intervals (0.05,0.06) and (0.06,0.07), both of which have positive $\widehat{p}_j$, and we add the intervals $(-\infty,0.05)$ and $(0.07,+\infty)$, to the left and to the right, respectively. We then compute the sums of the $\widetilde{p}_j$ in (A3), in each of these four intervals. Lastly, we use these cumulative probabilities to estimate μ_g and σ_g by OLS. Since, in the fourth interval, the cumulative probability is equal to 1, in this example we only rely on three independent linear restrictions to estimate μ_g and σ_g.

Last two waves. For the 1995–1998 waves, we use the survey expectations questions to estimate the mean μ and variance σ^2 of log income beliefs directly (since in these waves the questions are about income, not income growth). We interpret the answers as probabilities assigned to two bins (between the minimum and the mid-point, and between the mid-point and the maximum). As in the 1989–1991 waves, we add two additional bins, below the reported minimum and above the reported maximum, respectively, which amounts to be working with 4 bins in total. These additional bins have a positive but low probability $\widetilde{p}_j = \frac{1/2M}{1+2M}$, which might reflect that respondents interpret the minimum and maximum questions as asking them to report quantiles of their distributions (see Delavande, Giné, and McKenzie, 2011). In the 1995–1998 waves, the location and width of the bins come from individuals' responses, providing more information to capture beliefs, in particular beliefs with very small variance. For example, when the reported minimum and maximum coincide, the implied estimate of σ is equal to zero.

Descriptives and predictive power. In Table C3 we provide descriptive statistics about the beliefs that we estimate and the main variables in the consumption equation.

In Table C4 we assess the predictive power of these beliefs: we regress $\log(w_{i,t+1})$ in columns (1) to (4), and $\log(w_{i,t+1}) - \log(w_{it})$ in columns (5) to (8), as functions of the estimated mean

[2] We verified that our estimates of the log consumption function remain similar when using different bounds, and when excluding observations that assign all points to the first or last bin.

[3] We found that using bins with $\widehat{p}_j = 0$ tended to artificially increase the variance of estimated beliefs.

beliefs μ_{it} and other controls. In this table, we use log individual income as our dependent variable. The estimates suggest that individual beliefs predict future income, even conditional on current income.

Table C3: Descriptive statistics

	Cross-sectional sample				Panel sample			
	Obs	P25	Mean	P75	Obs	P25	Mean	P75
Log family consumption	7,796	9.78	10.05	10.31	1,646	9.78	10.07	10.33
Log family assets	7,496	10.03	11.04	12.18	1,587	10.33	11.21	12.28
Log family income	7,795	10.03	10.39	10.74	1,645	10.07	10.43	10.79
Log individual income	7,791	9.69	9.87	10.07	1,644	9.73	9.91	10.11
Mean expected log income	7,796	9.72	9.92	10.13	1,646	9.75	9.96	10.16
SD expected log income	7,796	0.005	0.015	0.017	1,646	0.005	0.015	0.017

Notes: Amounts are in 2010 euros. Descriptive statistics are weighted using the survey's weights. Individual income excludes property income and income from transfers. Individual-level variables (i.e., income and income expectations) corresponds to the household head.

Table C4: Predictive power of income beliefs

	$\log(w_{i,t+1})$				$\log(w_{i,t+1}) - \log(w_{it})$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean expected log income		0.596		0.367				
		(0.036)		(0.082)				
Mean expected change in log income						0.659		0.367
						(0.116)		(0.082)
Log individual income			0.566	0.239			-0.434	-0.394
			(0.041)	(0.083)			(0.041)	(0.038)
Sample	1989-1998	1989-1998	1989-1998	1989-1998	1989-1998	1989-1998	1989-1998	1989-1998
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N observations	2,994	2,994	2,994	2,994	2,994	2,994	2,994	2,994
R-squared	0.290	0.466	0.460	0.470	0.047	0.098	0.196	0.211

Notes: SHIW, 1989–1991 and 1995–1998. Regression for household heads. Controls include age and age squared, gender, education, indicator of spouse, marital status, family size, number of children 0-5, 6-13, 14-17 years old in the household, number of children outside the household, area, number of income earners in the household, and a wave indicator. Regression estimates are weighted using survey weights. Standard errors (shown in parenthesis) are clustered at the household level.

D Robustness checks

In this section of the appendix, we provide several robustness checks for the estimation of the consumption function.

In columns (1) and (2) in Table D1 we show the estimates are robust to relying on different distributional assumptions for beliefs: a discrete distribution for waves 1989–1991 (as in Pistaferri, 2001), and a triangular distribution for waves 1995-1998 (as in Kaufmann and Pistaferri, 2009). In columns (3) to (6) we show that estimates are robust to the value of M used for estimation (see (A3), where the baseline corresponds to $M = 100$). In columns (7) and (8) we also control for the spouse's beliefs about their own income, when available.[4] Results remain virtually unchanged, and spousal beliefs don't appear to play a major role in household consumption for this sample.

In Table D2 we estimate the consumption function, focusing on the specification with mean beliefs interacted with log current income, separately for waves 1989–1991 and 1995–1998.[5] The point estimates are different in the two samples, with a larger effect of beliefs in the 1995-1998 waves. However, in both cases beliefs play a significant role in household consumption.[6]

Lastly, in Table D3 we present estimates obtained under different approaches for dealing with assets. As mentioned in the main text, the estimates of current income and income beliefs are quite similar across specifications, although we see some quantitative differences, especially in the case of the IV specification in columns (5) and (6).

[4]When spousal beliefs are not available, we set the variable to zero and add binary indicators for missingness, distinguishing between spouses that are homemakers, employed, or other labor status. Note that only 32% and 17% of the 768 households in columns (7) and (8), respectively, are households where data on spousal beliefs are available in at least one or in both waves.

[5]In each pair of waves, we also control for other expectations questions available: inflation expectations in 1989–1991, and expectations about future employment in 1995–1998.

[6]Using the 1995–1998 waves, we also estimated the consumption function including unemployed household heads in the sample and controlling for beliefs about future employment, and found similar results. In the 1989–1991 waves expectations questions were not asked to the unemployed.

Table D1: Estimates of the log consumption function: robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean expected log income head	0.235	0.229	0.237	0.230	0.235	0.229	0.245	0.242
	(0.095)	(0.093)	(0.095)	(0.094)	(0.095)	(0.093)	(0.095)	(0.093)
(Mean expect. log income head)·(Log family income)		0.106		0.105		0.104		0.103
		(0.061)		(0.061)		(0.061)		(0.062)
Mean expected log income spouse							0.018	-0.022
							(0.054)	(0.064)
(Mean expect. log income spouse)·(Log family income)								0.011
								(0.009)
Log family income	0.438	0.438	0.438	0.438	0.439	0.439	0.428	0.439
	(0.091)	(0.090)	(0.090)	(0.089)	(0.089)	(0.089)	(0.091)	(0.091)
Log family assets	0.016	0.017	0.018	0.019	0.018	0.019	0.018	0.020
	(0.024)	(0.024)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distribution assumption	Disc - Triang	Disc - Triang	Log-normal	Log-normal	Log-normal	Log-normal	Log-normal	Log-normal
M draws			10	10	50	50	100	100
N observations	1,514	1,514	1,536	1,536	1,536	1,536	1,536	1,536
N households	757	757	768	768	768	768	768	768
R-squared	0.26	0.26	0.26	0.26	0.26	0.26	0.26	0.26
Pvalue F beliefs head	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.01
Pvalue F beliefs spouse							0.74	0.45
Pvalue F beliefs head and spouse							0.04	0.04

Notes: SHIW, regression for household heads. In columns (1) and (2) we assume a different distribution of beliefs (discrete distribution in waves 1989–1991 and triangular distribution in waves 1995–1998). In columns (3) to (6) we vary the number M of draws used in estimation. In columns (7) and (8), we add spouse's beliefs (for spouses that are employees and have beliefs questions, and 0 for everyone else). The expectations variables and log family income are centered around the weighted average in the sample. Controls include age and age squared, existence of a spouse, marital status, family size, number of children 0-5, 6-13, 14-17 years old in the household, number of children outside the household, number of income earners in the household, and a wave indicator. In columns (7) and (8), we also control for a categorical variable indicating spousal situation (no spouse, spouse is homemaker, spouse is employee with beliefs questions, spouse is employee without beliefs questions, other). Regression estimates are weighted using survey weights. Standard errors (shown in parenthesis) are clustered at the household level.

Table D2: Estimates of the log consumption function by wave

	(1)	(2)	(3)	(4)	(5)	(6)
Mean expected log income	0.235	0.229	0.212	0.242	0.323	0.342
	(0.094)	(0.093)	(0.110)	(0.108)	(0.171)	(0.172)
(Mean expect. log income)·(Log family income)		0.104		0.113		-0.125
		(0.061)		(0.060)		(0.177)
Log family income	0.439	0.439	0.461	0.442	0.277	0.264
	(0.089)	(0.089)	(0.101)	(0.100)	(0.169)	(0.168)
Log family assets	0.018	0.019	0.046	0.048	-0.063	-0.060
	(0.023)	(0.023)	(0.027)	(0.026)	(0.039)	(0.039)
Sample	1989-1998	1989-1998	1989-1991	1989-1991	1995-1998	1995-1998
Household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N observations	1,536	1,536	962	962	512	512
N households	768	768	481	481	256	256
R-squared	0.26	0.26	0.35	0.37	0.16	0.17
Pvalue F beliefs	0.01	0.02	0.05	0.03	0.06	0.14

Notes: SHIW, regression for household heads. The expectations variables and log family income are centered around the weighted average in the sample. Controls include age and age squared, existence of a spouse, marital status, family size, number of children 0-5, 6-13, 14-17 years old in the household, number of children outside the household, number of income earners in the household, and a wave indicator. When available, we also control for other expectations variables: columns (3) and (4) also control for mean expected inflation, and columns (5) and (6) also control for the beliefs about the probability of being employed next year. Regression estimates are weighted using survey weights. Standard errors (shown in parenthesis) are clustered at the household level.

Table D3: Estimates of the log consumption function: robustness to assets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean expected log income	0.235	0.229	0.245	0.238	0.167	0.159	0.223	0.216
	(0.094)	(0.093)	(0.097)	(0.095)	(0.107)	(0.106)	(0.096)	(0.095)
(Mean expect. log income)·(Log family income)		0.104		0.095		0.093		0.102
		(0.061)		(0.061)		(0.062)		(0.060)
Log family income	0.439	0.439	0.410	0.413	0.642	0.648	0.475	0.476
	(0.089)	(0.089)	(0.097)	(0.097)	(0.144)	(0.144)	(0.097)	(0.096)
Log family assets	0.018	0.019	0.033	0.032	-0.084	-0.087		
	(0.023)	(0.023)	(0.032)	(0.032)	(0.055)	(0.054)		
(Log family assets)^2			0.007	0.006				
			(0.006)	(0.006)				
Household fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IV	No	No	No	No	Yes	Yes	No	No
N observations	1,536	1,536	1,536	1,536	1,536	1,536	1,536	1,536
N households	768	768	768	768	768	768	768	768
R-squared	0.26	0.26	0.26	0.26	.	.	0.26	0.26
Pvalue F beliefs	0.01	0.02	0.01	0.02	0.12	0.13	0.02	0.02
Pvalue first stage					0.00	0.00		

Notes: SHIW, regression for household heads. In columns (5) and (6) we instrument the difference of log family assets by first-period assets and income. The expectations variables and log family income are centered around the weighted average in the sample. Controls include age and age squared, existence of a spouse, marital status, family size, number of children 0-5, 6-13, 14-17 years old in the household, number of children outside the household, number of income earners in the household, and a wave indicator. Regression estimates are weighted using survey weights. Standard errors (shown in parenthesis) are clustered at the household level.

E Measurement error

To assess the extent of measurement error in the beliefs responses, we use the 1989–1991 waves. In our baseline specification, we estimate the mean and variance of beliefs by relying on a model that assumes individuals draw $M = 100$ different scenarios from their underlying beliefs to answer the expectations questions (see Subsection C.2 of this appendix). This choice is motivated by the format of the questions, where respondents are asked to distribute 100 points among the bins.

However, this model may not provide a good approximation to the response process of individuals when answering the questions in the SHIW. In fact, it is possible that respondents are only able to imagine a smaller number $M < 100$ of "income growth scenarios", corresponding to events that they expect might happen in the next year, such as a promotion or a demotion, a job change, etc. To provide empirical support to this possibility, we predict, for each respondent, the number of bins with a non-zero empirical frequency under the model for various values of M, see Table E1. We find that taking $M = 100$ implies that, on average, respondents should report 3.6 non-empty bins. In contrast, in the data this number is only 1.7. Besides, the table shows that taking smaller values of M provides a better approximation to the distribution of the number of non-empty bins across individuals.

With this motivation, in this section of the appendix we entertain an alternative parametric model for the responses, where individuals draw $M < 100$ values from a $\mathcal{N}(\mu_g, \sigma_g^2)$, and distribute those among the bins.[7] Given this model, we propose a correction for measurement error and apply it to revisit our baseline estimates of the consumption function (see Table 2). Our approach is based on a "small-σ" approximation (e.g., Evdokimov and Zeleneev, 2022). Given that, for a given M value, the model of measurement error is parametric, the correction can be implemented using a simple parametric bootstrap method, which we now describe.[8]

We consider the specification of the consumption function in column (3) of Table D2, which only accounts for mean beliefs. We draw $S = 1,000$ samples where, for each respondent, we draw M observations from a $\mathcal{N}(\widehat{\mu}_g, \widehat{\sigma}_g^2)$, for $\widehat{\mu}_g$ and $\widehat{\sigma}_g^2$ our original estimates of μ_g and σ_g^2, respectively. This gives us simulated responses $\widehat{p}_j^{(s)}$, for each sample s, from which we estimate

[7] In the model of measurement error that we propose, M is constant across individuals. An alternative model would let M_i vary across individuals. Manski and Molinari (2010) exploit repeated responses by the same individual to infer individual types of measurement error in responses.

[8] Since the measurement error model is parametric, one could alternatively rely on an exact approach for deconvolving the measurement error, without the need for an approximation. An advantage of the specific approach that we implement here is its simplicity.

μ_g and σ_g and, based on those, the coefficients of the consumption function, exactly in the same way as we did to obtain the estimates in Table D2.[9] Let $\widehat{\beta}^{(s)}$ denote the estimated coefficients in this last regression. We then construct the bootstrapped bias-corrected counterpart to the original coefficients $\widehat{\beta}^{\text{OLS}}$ as

$$\widehat{\beta}^{\text{BC}} = 2\widehat{\beta}^{\text{OLS}} - \frac{1}{S}\sum_{s=1}^{S}\widehat{\beta}^{(s)}.$$

We repeat this exercise for values of M between 1 and 100.

In Figure E1 we report the bias-corrected estimator $\widehat{\beta}^{\text{BC}}$ for two of the regression parameters: the coefficient of the mean income beliefs, and the coefficient of current log income. We report the results for different values of M. The figure shows that the results are fairly robust to this form of measurement error, with $\widehat{\beta}^{\text{BC}}$ and $\widehat{\beta}^{\text{OLS}}$ being close to each other irrespective of M. In addition, the variability induced by this form of measurement error, as captured by the dashed lines in the figure, appears moderate.

[9] In particular, we still consider an uninformative prior with 100 trials.

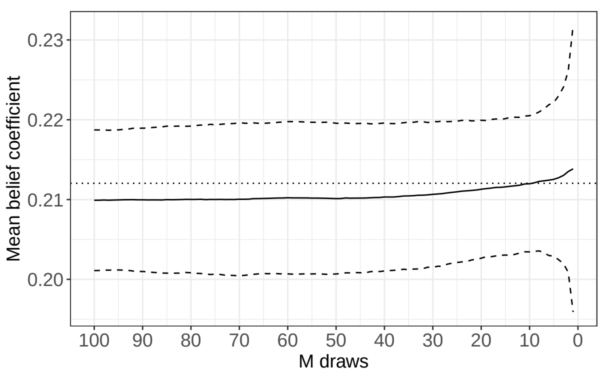
Table E1: Predicted distribution of number of bins by number of draws M

	Number of bins with non-zero frequencies												
	1	2	3	4	5	6	7	8	9	10	11	12	Mean
Data	0.59	0.24	0.09	0.05	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	1.75
$M=1$	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
$M=2$	0.68	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.32
$M=3$	0.57	0.35	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.51
$M=4$	0.50	0.36	0.12	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.66
$M=5$	0.45	0.37	0.14	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.78
$M=6$	0.42	0.37	0.15	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.88
$M=7$	0.39	0.37	0.16	0.06	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.96
$M=8$	0.36	0.38	0.16	0.06	0.02	0.01	0.00	0.00	0.00	0.00	0.00	0.00	2.03
$M=9$	0.34	0.38	0.17	0.07	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	2.10
$M=10$	0.32	0.39	0.18	0.07	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	2.16
$M=20$	0.17	0.41	0.24	0.09	0.05	0.02	0.01	0.00	0.00	0.00	0.00	0.00	2.59
$M=30$	0.09	0.39	0.30	0.11	0.06	0.03	0.01	0.01	0.00	0.00	0.00	0.00	2.87
$M=40$	0.05	0.36	0.34	0.13	0.06	0.03	0.02	0.01	0.00	0.00	0.00	0.00	3.07
$M=50$	0.03	0.31	0.38	0.14	0.07	0.04	0.02	0.01	0.00	0.00	0.00	0.00	3.22
$M=60$	0.01	0.28	0.41	0.15	0.07	0.04	0.02	0.01	0.00	0.00	0.00	0.00	3.33
$M=70$	0.01	0.24	0.43	0.16	0.08	0.04	0.02	0.01	0.00	0.00	0.00	0.00	3.42
$M=80$	0.00	0.21	0.45	0.16	0.08	0.04	0.02	0.01	0.00	0.00	0.00	0.00	3.49
$M=90$	0.00	0.19	0.46	0.16	0.09	0.04	0.03	0.01	0.01	0.00	0.00	0.00	3.55
$M=100$	0.00	0.16	0.48	0.17	0.09	0.05	0.03	0.01	0.01	0.00	0.00	0.00	3.61

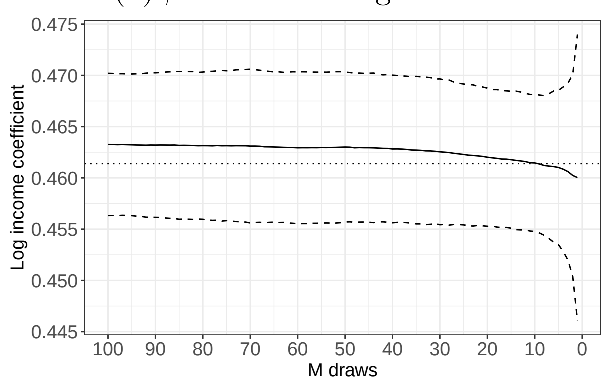
Notes: SHIW, 1989–1991, sample from column (3) in Table D2. Each row reports the simulated distribution of the number of non-empty bins in data simulated from a measurement error model with M draws, averaged across observations and $S=1,000$ simulations. Unweighted results.

Figure E1: Bias-corrected coefficients of mean beliefs and log income

(a) β for mean income beliefs



(b) β for current log income



Notes: SHIW, 1989–1991, sample from column (3) in Table D2. The horizontal dotted lines show the corresponding elements of $\widehat{\beta}^{\mathrm{OLS}}$ *from column (3) in Table D2. The solid lines show* $\widehat{\beta}^{\mathrm{BC}}$*, and the dashed lines add a band of plus or minus twice the standard deviation of* $\widehat{\beta}^{(s)}$ *across simulations.* $1,000$ *simulations.*

F Tax counterfactuals: details about estimation

In this section of the appendix we detail the calculations of tax counterfactuals and present additional empirical estimates.

F.1 Tax schedule

We assume the tax schedule takes the parametric form $T(\widetilde{w}_r) = \widetilde{w}_r - \lambda \widetilde{w}_r^{1-\tau}$, where $\widetilde{w}_r$ denotes gross income in multiples of its population average, as in Benabou (2002). This parametric form can be re-written as a similar function that depends on gross income $\widetilde{w}$, with the same parameter τ but a different parameter $\widetilde{\lambda}$.[10] For the baseline level of the tax, we rely on the estimates obtained by Holter, Krueger, and Stepanchuk (2019) for Italy, averaged over family composition characteristics in our sample: $\lambda_0 = 0.94$ and $\tau_0 = 0.196$.

Let λ_1 and τ_1 denote the parameters defining the tax schedule under a counterfactual scenario. We assume the tax schedule applies to gross family income, and that each individual pays taxes proportionally to their contribution in the family, r_{it}, a proportion we assume does not change in counterfactual scenarios. Let x_{1it} denote log family income and (μ_{1it},σ^2_{1it}) denote the parameters of income beliefs under a counterfactual scenario. Let (x_{0it}, μ_{0it}, σ^2_{0it}) denote their baseline values, observed in sample. We obtain

$$\begin{aligned}
\mu_{1it} - \mu_{0it} &= \left[\log(\widetilde{\lambda}_1) - \frac{(1-\tau_1)}{(1-\tau_0)}\log(\widetilde{\lambda}_0)\right] + \frac{(\tau_0 - \tau_1)}{(1-\tau_0)}\mu_{0it} + \log(r_{it})\frac{\tau_1 - \tau_0}{1-\tau_0}, \\
\sigma^2_{1it} - \sigma^2_{0it} &= \sigma^2_{0it}\left[\frac{(1-\tau_1)^2}{(1-\tau_0)^2} - 1\right], \\
x_{1it} - x_{0it} &= \log(\widetilde{\lambda}_1) - \log(\widetilde{\lambda}_0) + \left[\frac{x_{0it} - \log(\widetilde{\lambda}_0)}{1-\tau_0}\right](\tau_0 - \tau_1).
\end{aligned}$$

Given a counterfactual tax schedule (λ_1, τ_1), we can use these values to compute average partial effects.

We consider three counterfactual scenarios. In the *transitory tax* increase and *permanent tax* increase counterfactuals, we set $\lambda_1 = \lambda_0 - 0.1$ and $\tau_1 = \tau_0$. In the *regressivity* counterfactual, we set $\tau_1 = 0.142$, the progressivity parameters of the tax system in France according to Holter, Krueger, and Stepanchuk (2019), and set λ_1 such that the tax change is revenue neutral.[11]

[10] Specifically, $\widetilde{\lambda} = \lambda K^\tau$, for K the average gross income in the population.

[11] Assuming that family gross income is log-normally distributed with parameters $\mu_{\widetilde{w}}$ and $\sigma^2_{\widetilde{w}}$, a change in the

F.2 Double Lasso estimation

In this subsection we describe how we estimate the consumption function using the double Lasso method introduced by Belloni, Chernozhukov, and Hansen (2014). Consider the equation,

$$y_{it} = a'\Psi(s_{it}) + \beta_k k_{it} + \alpha_i + \varepsilon_{it}, \tag{A5}$$

where $\Psi(s_{it})$ includes polynomial functions of the main covariates (age, log income, log assets, and the income beliefs' means and variances), and k_{it} includes the other demographic controls. Under this specification, an average partial effect corresponding to a counterfactual of interest is given by

$$a'\left(\frac{1}{nT}\sum_{i,t}\left(\Psi(\widetilde{s}_{it}) - \Psi(s_{it})\right)\right)$$

where s_{it} are the main covariates under the baseline, and $\widetilde{s}_{it}$ are the main covariates under the counterfactual.

Letting

$$v = \frac{1}{nT}\sum_{i,t}\left(\Psi(\widetilde{s}_{it}) - \Psi(s_{it})\right),$$

we first reparameterize the polynomials so that the average partial effect corresponds simply to the coefficient associated with the first regressor. To that end, we construct an invertible matrix A whose first column is equal to v.[12] Then, we rewrite (A5) using the reparameterized polynomials $\widetilde{\Psi}(s_{it}) = A^{-1}\Psi(s_{it})$, and obtain

$$y_{it} = (A'a)'\,\widetilde{\Psi}(s_{it}) + \beta_k k_{it} + \alpha_i + \varepsilon_{it}. \tag{A6}$$

Note that the coefficient of the first covariate in (A6) is equal to $a'v$, which is the average partial effect of interest.

To estimate $a'v$, we apply the double Lasso estimator to (A6). To account for household fixed effects, we take first differences. We always include (i.e., we do not penalize) the following regressors: the first order polynomials (age, log income, log assets, and the beliefs' means and

parameters of the tax system is revenue neutral if

$$\log(\widetilde{\lambda}_1) - \log(\widetilde{\lambda}_0) = \frac{1}{2}\sigma^2_{\widetilde{w}}\left[(1-\tau_0)^2 - (1-\tau_1)^2\right] + \mu_{\widetilde{w}}(\tau_1 - \tau_0)$$

Furthermore, $\mu_{\widetilde{w}} = (\mu_x - log(\tilde{\lambda}_0))/(1-\tau_0)$ and $\sigma_{\widetilde{w}} = \sigma_x/(1-\tau_0)$, where μ_x and σ^2_x are the mean and variance of the log of disposable family income, which we estimate from the SHIW.

[12] For example, we set $A = [v \quad \iota_2 .. \quad \iota_L]$, where ι_ℓ are the canonical vectors in $\mathbb{R}^L$ and $L = \dim\Psi$, provided such a matrix A is invertible.

variances), as well as the variables in k_{it} (existence of a spouse, marital status, family size, number of children 0-5, 6-13, 14-17 years old in the household, number of children outside the household, number of income earners in the household, and a wave indicator).

The double Lasso method is implemented in two steps. In a first step, we apply the Lasso to regress the first element in $\widetilde{\Psi}(s_{it})$ on its second to last elements and k_{it}, in first differences. In the second step, we apply the Lasso to regress the y_{it} on the second to last elements of $\widetilde{\Psi}(s_{it})$ and k_{it}, in first differences. In both steps, we choose the penalty parameters by 10-fold cross-validation (Chetverikov, Liao, and Chernozhukov, 2021). Lastly, we regress y_{it} on the first element in $\widetilde{\Psi}(s_{it})$ and all the controls selected in the two Lasso steps, again in first differences. We account for the estimation uncertainty (in particular, the fact that v is estimated) by computing bootstrapped standard errors.

F.3 Empirical estimates

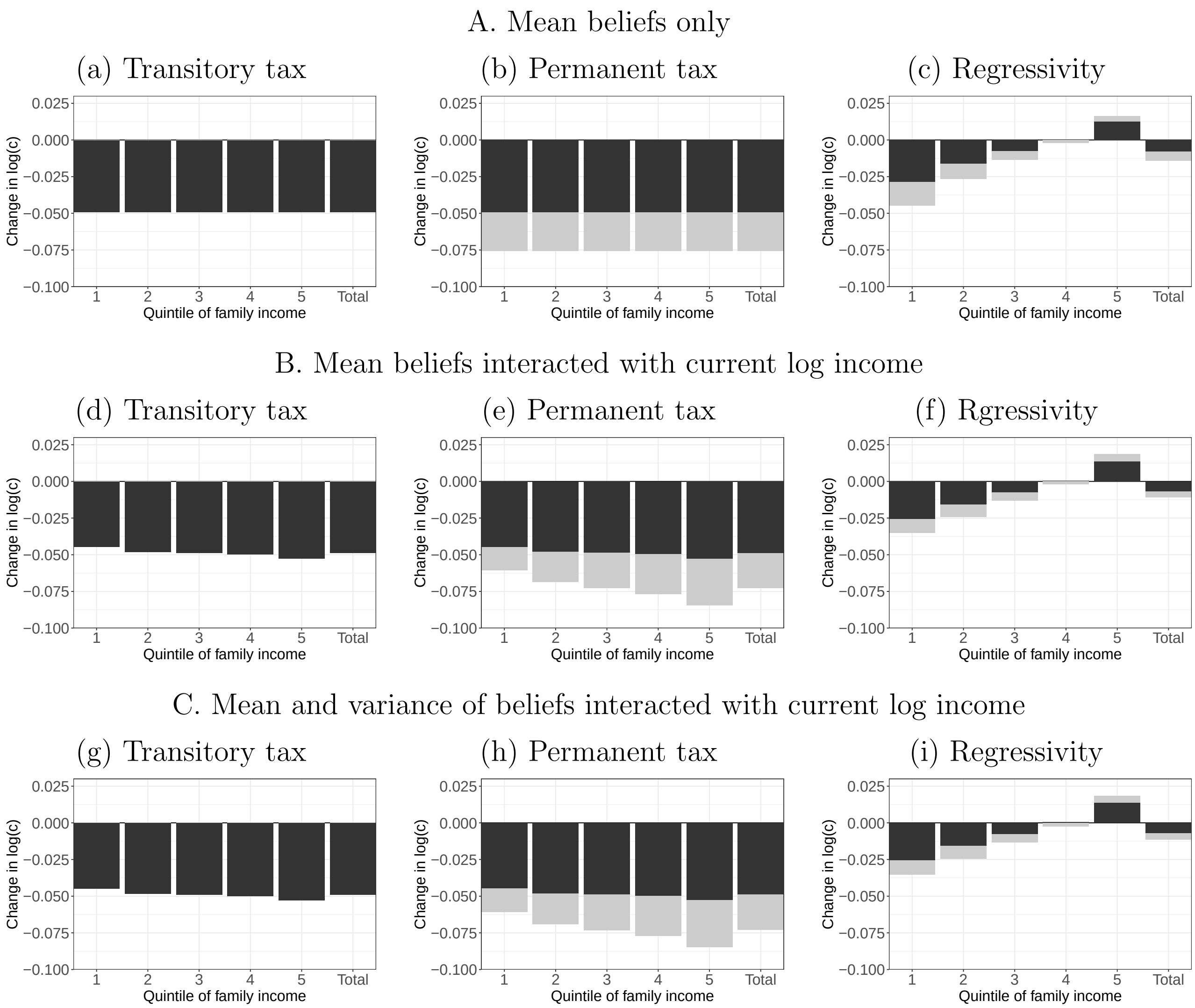
In Table F1 we report average partial effects based on OLS estimates of the consumption function, and in Table F2 we report average partial effects based on the double Lasso. We show these in graphical form in Figures F1 and F2, respectively. Overall, the results are quite consistent across specifications.

Table F1: Average partial effects estimates (OLS)

Quintile	*Transitory tax* counterfactual			*Permanent tax* counterfactual			*Regressivity* counterfactual		
	CAPE	DAPE	TAPE	CAPE	DAPE	TAPE	CAPE	DAPE	TAPE
A. Mean beliefs only									
1	-0.0493	0.0000	-0.0493	-0.0493	-0.0264	-0.0757	-0.0288	-0.0161	-0.0449
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0060)	(0.0064)	(0.0052)
2	-0.0493	0.0000	-0.0493	-0.0493	-0.0264	-0.0757	-0.0162	-0.0105	-0.0267
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0034)	(0.0042)	(0.0033)
3	-0.0493	0.0000	-0.0493	-0.0493	-0.0264	-0.0757	-0.0076	-0.0058	-0.0134
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0016)	(0.0023)	(0.0018)
4	-0.0493	0.0000	-0.0493	-0.0493	-0.0264	-0.0757	0.0004	-0.0020	-0.0015
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0004)	(0.0008)	(0.0009)
5	-0.0493	0.0000	-0.0493	-0.0493	-0.0264	-0.0757	0.0126	0.0036	0.0162
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0026)	(0.0015)	(0.0021)
Total	-0.0493	0.0000	-0.0493	-0.0493	-0.0264	-0.0757	-0.0079	-0.0061	-0.0141
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0017)	(0.0024)	(0.0019)
B. Mean beliefs interacted with current log income									
1	-0.0448	0.0000	-0.0448	-0.0448	-0.0157	-0.0605	-0.0257	-0.0094	-0.0351
	(0.0105)	(0.0000)	(0.0105)	(0.0105)	(0.0117)	(0.0117)	(0.0062)	(0.0072)	(0.0073)
2	-0.0481	0.0000	-0.0481	-0.0481	-0.0206	-0.0687	-0.0158	-0.0085	-0.0243
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0107)	(0.0091)	(0.0034)	(0.0042)	(0.0034)
3	-0.0488	0.0000	-0.0488	-0.0488	-0.0240	-0.0728	-0.0075	-0.0055	-0.0130
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0104)	(0.0086)	(0.0016)	(0.0023)	(0.0018)
4	-0.0496	0.0000	-0.0496	-0.0496	-0.0271	-0.0768	0.0005	-0.0021	-0.0016
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0087)	(0.0004)	(0.0008)	(0.0009)
5	-0.0527	0.0000	-0.0527	-0.0527	-0.0319	-0.0845	0.0138	0.0049	0.0187
	(0.0104)	(0.0000)	(0.0104)	(0.0104)	(0.0113)	(0.0107)	(0.0027)	(0.0017)	(0.0026)
Total	-0.0488	0.0000	-0.0488	-0.0488	-0.0239	-0.0727	-0.0070	-0.0041	-0.0111
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0104)	(0.0086)	(0.0018)	(0.0026)	(0.0024)
C. Mean and variance of beliefs interacted with current log income									
1	-0.0449	0.0000	-0.0449	-0.0449	-0.0160	-0.0608	-0.0257	-0.0097	-0.0355
	(0.0105)	(0.0000)	(0.0105)	(0.0105)	(0.0119)	(0.0118)	(0.0063)	(0.0077)	(0.0079)
2	-0.0482	0.0000	-0.0482	-0.0482	-0.0209	-0.0691	-0.0158	-0.0088	-0.0246
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0108)	(0.0091)	(0.0034)	(0.0043)	(0.0035)
3	-0.0489	0.0000	-0.0489	-0.0489	-0.0242	-0.0731	-0.0075	-0.0057	-0.0132
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0016)	(0.0024)	(0.0019)
4	-0.0498	0.0000	-0.0498	-0.0498	-0.0274	-0.0771	0.0005	-0.0023	-0.0018
	(0.0103)	(0.0000)	(0.0103)	(0.0103)	(0.0106)	(0.0088)	(0.0004)	(0.0009)	(0.0011)
5	-0.0528	0.0000	-0.0528	-0.0528	-0.0321	-0.0849	0.0138	0.0047	0.0185
	(0.0105)	(0.0000)	(0.0105)	(0.0105)	(0.0114)	(0.0108)	(0.0028)	(0.0018)	(0.0027)
Total	-0.0489	0.0000	-0.0489	-0.0489	-0.0241	-0.0730	-0.0070	-0.0044	-0.0113
	(0.0102)	(0.0000)	(0.0102)	(0.0102)	(0.0105)	(0.0086)	(0.0018)	(0.0027)	(0.0026)

Notes: SHIW, 1989–1991 and 1995–1998, cross-sectional sample. The top panel is based on column (2) in Table 2, the middle panel on column (4), and the bottom panel on column (5). Standard errors are based on 1,000 bootstrap replications.

Figure F1: Average partial effects estimates (OLS)



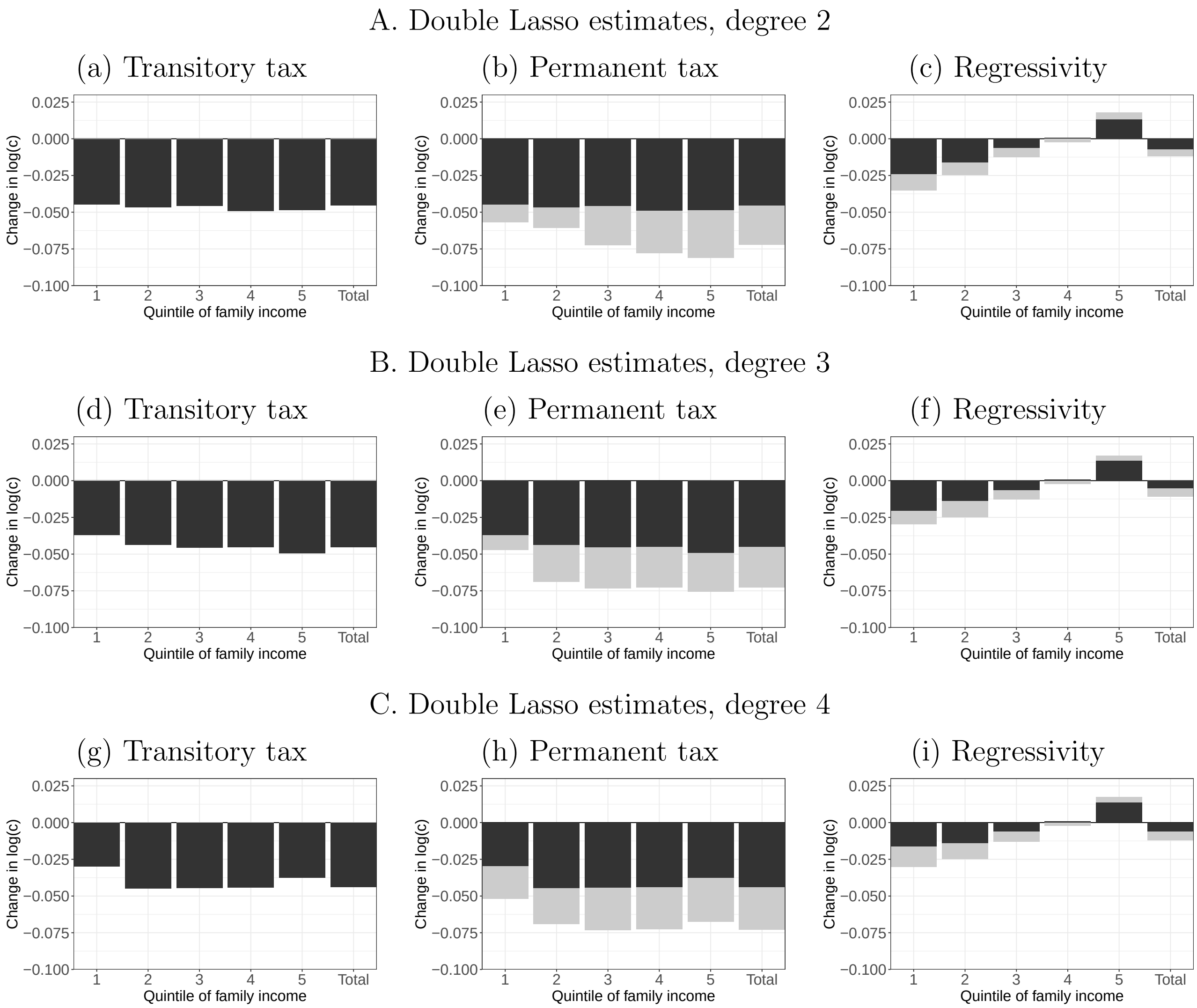
Notes: SHIW, 1989–1991 and 1995–1998, cross-sectional sample. Black bars correspond to contemporaneous APE and grey bars correspond to dynamic APE. Total APE are the sums of CAPE and DAPE. The top panel is based on column (2) in Table 2, the middle panel on column (4), and the bottom panel on column (5).

Table F2: Average partial effects estimates (Lasso), in progress

Quintile	*Transitory tax* counterfactual			*Permanent tax* counterfactual			*Regressivity* counterfactual		
	CAPE	DAPE	TAPE	CAPE	DAPE	TAPE	CAPE	DAPE	TAPE
A. Double Lasso estimates, degree 2									
1	-0.0447	0.0000	-0.0447	-0.0447	-0.0120	-0.0567	-0.0240	-0.0111	-0.0351
	(0.0183)	(0.0000)	(0.0183)	(0.0183)	(0.0182)	(0.0147)	(0.0107)	(0.0111)	(0.0094)
2	-0.0467	0.0000	-0.0467	-0.0467	-0.0140	-0.0607	-0.0161	-0.0083	-0.0245
	(0.0143)	(0.0000)	(0.0143)	(0.0143)	(0.0147)	(0.0113)	(0.0043)	(0.0052)	(0.0041)
3	-0.0459	0.0000	-0.0459	-0.0459	-0.0266	-0.0725	-0.0065	-0.0061	-0.0126
	(0.0119)	(0.0000)	(0.0119)	(0.0119)	(0.0114)	(0.0086)	(0.0017)	(0.0025)	(0.0019)
4	-0.0491	0.0000	-0.0491	-0.0491	-0.0289	-0.0780	0.0008	-0.0024	-0.0015
	(0.0130)	(0.0000)	(0.0130)	(0.0130)	(0.0107)	(0.0106)	(0.0004)	(0.0013)	(0.0014)
5	-0.0486	0.0000	-0.0486	-0.0486	-0.0325	-0.0811	0.0132	0.0045	0.0177
	(0.0180)	(0.0000)	(0.0180)	(0.0180)	(0.0118)	(0.0149)	(0.0056)	(0.0025)	(0.0048)
Total	-0.0455	0.0000	-0.0455	-0.0455	-0.0265	-0.0720	-0.0072	-0.0047	-0.0119
	(0.0120)	(0.0000)	(0.0120)	(0.0120)	(0.0115)	(0.0084)	(0.0034)	(0.0036)	(0.0033)
B. Double Lasso estimates, degree 3									
C. Double Lasso estimates, degree 4									

Notes: SHIW, 1989–1991 and 1995–1998, cross-sectional sample. The top panel is based on polynomials of degree 2, the middle panel on polynomials of degree 3, and the bottom panel on polynomials of degree 4. Standard errors are based on 1,000 bootstrap replications.

Figure F2: Average partial effects estimates (Lasso)



Notes: SHIW, 1989–1991 and 1995–1998, cross-sectional sample. Black bars correspond to contemporaneous APE and grey bars correspond to dynamic APE. Total APE are the sums of CAPE and DAPE. Double Lasso estimates. The top panel is based on polynomials of degree 2, the middle panel on polynomials of degree 3, and the bottom panel on polynomials of degree 4.

G Extensions

In this section we further discuss two extensions of our approach that we introduced in Section 7.

G.1 State-contingent beliefs

The framework can be generalized to allow for endogenous and exogenous states in x_{it}, and for contingent beliefs about them. To see this, suppose for simplicity that actions y_{it} belong to a finite set $\mathcal{Y}$ with n_y elements. In this case, one can define $\pi_{it} = \{\pi_{it}(\cdot; y) : y \in \mathcal{Y}\}$ to be a set of n_y conditional distributions, which satisfies, for all x, π, y,

$$(x_{i,t+1} \,|\, x_{it} = x, \pi_{it} = \pi, y_{it} = y) \sim \pi(\cdot; y).$$

A difference with the framework of Section 3 is that in this case actions depend on the n_y distributions $\pi_{it}(\cdot; y)$.

G.2 Beliefs over longer horizons

A key feature of the framework is that, while beliefs about next period's state variables change in the data and counterfactual, the belief updating rule ρ_i is constant in sample and invariant to the counterfactual change. This assumption can be relaxed by introducing beliefs over multiple horizons.

To describe such an approach, let us replace part 2 in Assumption 1 by the following, for some $S \geqslant 1$:

$$(x_{i,t+S}, ..., x_{i,t+1} \,|\, x_{it} = x, \pi_{it} = \pi) \sim \pi, \quad \text{for all } x, \pi. \tag{A7}$$

In this case, (6) becomes

$$\begin{aligned} V_i(x_t, z_t, \pi_t) = \max_{y_t} \Big\{ & u_i(y_t, x_t, z_t) \\ & + \beta_i \iint V_i(x_{t+1}, \gamma_i(z_t, x_t, y_t), \pi_{t+1}) \pi_t^{(1)}(x_{t+1}) \rho_i(\pi_{t+1}; x_{t+1}, \pi_t, x_t) dx_{t+1} d\pi_{t+1} \Big\}, \end{aligned}$$

where $\pi_t^{(1)}$ denotes the marginal of π_t corresponding to period-$t+1$ outcomes. Hence, equation (7) is satisfied for the π_{it} defined in (A7).