Persistent heterogeneity in cognitive skills and overconfidence, liquidity constraints, and fiscal policy

Oliver Pfäuti* Fabian Seyrich† Jonathan Zinman‡

First Version: March 2023
This version: December 2023
Link to most recent version

Abstract

Heterogeneity in households' savings behavior and financial situations significantly affects macroeconomic fluctuations and policy. Using micro data on households' financial situations and cognitive skills, we uncover a systematic relationship between these two. Cognitively-less skilled households are more likely to be hand-to-mouth and overconfident about their skills. Households that are overconfident about their skills are also likely to be overly optimistic about their future financial situations. Given these empirical insights, we develop a Heterogeneous Agent New Keynesian model with heterogeneity in cognitive skills and overconfidence. The model accounts for our empirical findings and jointly matches the average marginal propensity to consume and the average wealth level even when all wealth is liquid. Allowing for heterogeneity in cognitive skills and overconfidence has important normative and positive implications for fiscal policy: the optimal government debt level is substantially lower, and—for a given average MPC—transfers to low-income households are less stimulating.

*University of Bern, Department of Economics, pfaeuti.oliver@gmail.com

†Berlin School of Economics, DIW Berlin, and Freie Universität Berlin, fabian.seyrich@gmail.com

‡Dartmouth College, Department of Economics, jzinman@dartmouth.edu.

An earlier version of this paper was titled "Heterogeneity in what? Cognitive Skills, Beliefs and the Liquid Wealth Distribution". We thank Eduardo Dávila, Greg Kaplan, Kurt Mitman, Peter Maxted, Matthew Rognlie, Johannes Stroebel, Hannes Twieling, Gianluca Violante, Michael Weber, and conference participants at the Bank of Finland, the 13th ifo Conference on Macroeconomics and Survey Data, EEA/ESEM 2023, and 5th Behavioral Macroeconomics Workshop for helpful comments and suggestions. Oliver Pfäuti gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project C02) and the Stiftung Geld & Währung. Fabian Seyrich gratefully acknowledges financial support by the Leibniz Association through the project "Distributional effects of macroeconomic policies in Europe".

1 Introduction

Heterogeneity in households' savings behavior and in their financial situations has significant implications for macroeconomic fluctuations and policy design.[1] Yet it remains standard practice in macro modeling to assume ex-ante identical households and account for heterogeneity only in shock realizations, abstracting from fundamental and persistent dimensions of heterogeneity that may shape households' savings behavior.[2] One such dimension of systematic heterogeneity across households is in cognitive skills, which has been linked empirically to differences in economic growth across space and time (Hanushek and Woessmann (2008)), consumers' roles in driving economic fluctuations (D'Acunto et al. (2019, 2023a,b)), financial mistakes (Agarwal and Mazumder (2013)), and strong negative relationships between behavioral biases and income (e.g., Stango and Zinman (2023), Chapman et al. (2023)).

Building on that work, we empirically document a systematic relationship between households' cognitive skills and their savings behavior, including precautionary savings and hand-to-mouth (HTM) status. Guided by this empirical evidence, we add persistent heterogeneity regarding cognitive skills to an otherwise state-of-the-art heterogeneous agent New Keynesian (HANK) model. Our model does well in jointly matching our empirical facts, the wealth distribution and the average marginal propensity to consume (MPC). We further show that accounting for the systematic relationship between households' cognitive skills and their savings behavior has important implications for fiscal policy, both positively and normatively.

We start by using micro data on U.S. consumers from the American Life Panel to develop several new facts about how cognitive skills, beliefs, and household financial situations are related. The likelihood of being persistently HtM, measured in various ways, decreases sharply with cognitively skills. But allowing for cognitive skills heterogeneity alone is unlikely to help fit the macro data, as we later formalize, because permanently low-productivity households will still tend to save their way out of HtM status if they are classically rational. This motivates considering how households perceive their cognitive skills as well.

We show that persistent overestimation of one's own skills is prevalent and correlates strongly (negatively) with cognitive skills. Households that are overconfident about their skills are also about 1.5 times as likely than their well-calibrated counterparts to be persistently overly-optimistic about their future financial situations (measured using standard consumer sentiment questions). This suggests that lower-skilled consumers may be HtM at least in part because of their overconfidence. Consistent with this conjecture, we find strong positive correlations between persistent

[1]See, e.g., Werning (2015), Kaplan et al. (2018), Auclert (2019), Bayer et al. (2022), Luetticke (2021), Hagedorn et al. (2019), Patterson (2023), Almgren et al. (2022), Holm et al. (2021) on shock transmission and policy efficacy, and Dávila and Schaab (2023), McKay and Wolf (2023), Bhandari et al. (2021), Bilbiie (2021), Smirnov (2022), Yang (2022) on optimal policy design.

[2]Important exceptions include models allowing for heterogeneity in presumed classical preferences. Recent examples include Auclert et al. (2020), Aguiar et al. (2021), Kaplan and Violante (2022) and Kekre and Lenel (2022), which we discuss below.

overconfidence and our persistent HtM measures.

Equipped with our empirical findings, we add persistent heterogeneity in cognitive skills and perceptions thereof to a standard HANK model with otherwise ex-ante identical households, incomplete markets, idiosyncratic risk, borrowing constraints, a nominal rigidity in the form of sticky wages, and one asset in the form of liquid government bonds. We model skills as labor market productivity that is subject to idiosyncratic shocks, and overconfidence as overweighting the probability of reaching a better productivity state and underweighting the probability of reaching a worse state.[3] Motivated by our micro data on the prevalence of persistent overconfidence and the strong correlation between cognitive skills and overconfidence, and in the interest of parsimony, we calibrate the model such that 62% of households are high-skilled with well-calibrated beliefs about future productivity while the remaining 38% are low-skilled and overconfident.[4] Apart from differences in average skill levels, this introduces one new parameter—the degree of overconfidence—which we discipline to match one of our key new empirical findings, namely that lower-skilled households are about 1.5 times as likely to be overly optimistic about their future financial situations.

In contrast to standard one-asset HANK models and to a HANK model with heterogeneity in skills but not in beliefs about them, our model fares well in jointly matching total wealth in the economy, high HtM prevalence, and an average quarterly marginal propensity to consume (MPC) in the consensus range of 15-25% (e.g., Jappelli and Pistaferri (2010) and Havranek and Sokolova (2020)). Our model achieves reconciliation because overconfident households underestimate their insurance needs and consequently perceive the price of the asset as too high to merit accumulating a sufficient buffer stock. Thus, even when the supply of assets is high, a large share of overconfident households choose to hold only a small buffer stock and often end up being HtM, consistent with our empirical findings. This holds even when all wealth is liquid and held in a single asset.[5] These results are driven by differences in overconfidence rather than by differences in skills: removing overconfidence from the model by imposing rational beliefs for all households, while retaining heterogeneous average skill levels, fails to match the average MPC and delivers very few HtM

[3]See also Caplin et al. (2023) on underestimation of income risk who find that subjective earnings risk is significantly smaller than earnings risk inferred from administrative data.

[4]Results do not depend on our calibration strategy that all overconfident households are low skilled, because differences in HtM status are mainly driven by differences in overconfidence whereas differences in average skill levels play virtually no role in explaining heterogeneity in HtM status. Additionally, allowing for underconfident households barely affects our results on HtM status, but it brings the model even closer to the empirical estimate of the top 10% wealth share.

[5]Our model also accounts well for several untargeted wealth inequality statistics. It produces more and empirically realistic inequality than its rational counterpart, better matching empirical wealth shares—e.g., of the top 10% or the bottom 50%. Further, our model does not suffer from the "missing middle" problem (Kaplan and Violante, 2022). The missing middle problem refers to the issue that a model's implied wealth distribution is too polarized compared to the data. One way to see this is that median wealth to mean annual earnings ratio is about a magnitude smaller in a standard one-asset HANK model than in the data. Although our model does produce some polarization, with most overconfident households having low wealth and mostly rational households among the wealthiest, there are enough rational households who experience several periods of relatively low productivity levels to populate the middle of the wealth distribution.

households. Existing one-asset HA(NK) models struggle to match the data because if the supply of assets is large enough to match the average wealth in the economy, the price of the asset is so low that almost all households accumulate a sufficient buffer stock to make the borrowing constraint nonbinding (Auclert et al. (2023), Kaplan and Violante (2022)). This makes HtM status counterfactually rare implying that most households have low MPCs. Consequently, these models produce an average MPC that is too low.

A standard practice for reconciling HANK models with the data is to introduce a second illiquid asset that can be adjusted only infrequently (Kaplan and Violante (2014), Kaplan et al. (2018), Bayer et al. (2019), Auclert et al. (2023)). A drawback to this approach is that it produces a liquidity premium that is arguably too high, as discussed in Kaplan and Violante (2022). We show that a two-asset version of our model can fit the data with a liquidity premium that is substantially lower, because overconfident households underestimating their individual income risk implies that they also underestimate the shadow value of future liquidity and thereby put downward pressure on the equilibrium liquidity premium.[6]

Our model thus requires only one additional parameter to overcome several shortcomings of existing models while simultaneously accounting for our new findings on skills, beliefs and financial situations. It also turns out that the mechanism that allows us to better match these key features of micro and macro data—lower-skilled households' undersaving due to overconfidence about their future financial situations—generates important implications for fiscal policy as well.

Normatively, the optimal government debt level is substantially lower than in a standard HANK model.[7] Higher public debt increases the self-insurance possibilities of households (Woodford (1990), Aiyagari and McGrattan (1998)), but this benefit is lower in our model because overconfident households place lower value on the insurance function of cheaper assets. Even for high public debt levels, the share of HtM households remains high in our model as the overconfident households do not save themselves out of being constrained. This is also reflected in the wealth share of the bottom 50%: while their wealth share increases strongly in the rational model when the amount of liquidity increases, it remains stubbornly low in our model with overconfident households.

Positively, allowing for heterogeneity in overconfidence has several implications for another important fiscal policy tool: transfers to low-income households. Targeted transfers are commonly deployed with both long-run insurance and short-run stabilization objectives in mind. Their effects depend on household precautionary savings behavior and MPCs, which are each strongly influenced by heterogeneity in cognitive skills and overconfidence.

[6]The rational two-asset model produces a risk premium of 8.2%, whereas empirical estimates are in the ballpark of 5%. Our two-asset model with overconfidence produces a risk premium of 2.2%. Accounting for aggregate risk would likely to push up the model-implied risk premia, bringing our model's risk premium closer to the data and pushing the rational model's even further away from it.

[7]This holds independently of whether households can only save in government bonds or also in illiquid productive capital.

Long-run insurance provision through targeted transfers is relatively effective in our model because overconfidence mutes crowd out of self-insurance in stationary equilibrium. In a rational model, transfers generate significant crowd out and increase the share of households being HtM and the real interest rate in steady-state. Those increases are dampened in our model because overconfident households do little precautionary saving to begin with, creating less scope for crowd out.

Short-run stabilization through income-targeted transfers is less effective in our model compared to a rational model once the rational model is calibrated to target the same average MPC due to two mechanisms.[8] First, the average MPC of transfer recipients is lower even when the overall average MPC is the same across models. The reason is that in our model, overconfidence, not income, is the key predictor of being HtM and having a high MPC. Our model thus produces many households with high MPCs that are not low-income. In the rational model, in contrast, low income is an almost perfect proxy for being HtM: Nearly all HtM households are low-income and therefore receive transfers. Using our micro data on HtM status, we show that this is at odds with the data. The rational model overpredicts the share of HtM households in the lowest income quartile whereas our model with overconfidence matches HtM shares along the income distribution quite well. The second mechanism is a weakening of the relaxed precautionary saving motive highlighted by Bayer et al. (2023a). Overconfident households do not decrease their precautionary savings in the presence of higher transfers like rational households do, because they have little precautionary saving to begin with (and underestimate its insurance value).[9] Our model thus highlights the importance of accounting for systematic differences between financially constrained and unconstrained households when it comes to stabilization policies.

Related literature. We contribute to four strands of literature. One considers how cognitive skills heterogeneity affects the macroeconomy. So far, this literature is largely empirical and focused on growth (Hanushek and Woessmann (2008)). D'Acunto et al. (2019, 2023a,b) bring cognitive skills heterogeneity to the empirical study of economic fluctuations, showing it plays key roles in how households form their inflation expectations and respond (or not) to information and incentives provided by policy interventions. Our contributions to this literature are twofold. Empirically, we link heterogeneity in cognitive skills to heterogeneity in current and forecasted financial situations (including HtM status). Theoretically, we build a model that captures the

[8] We take the empirical estimates of average MPCs as given here; as is well-understood, in a standard rational one-asset HANK model this requires artificially reducing the amount of total wealth (in our calculations, down to 0.6 of average annual income from the empirical standard of 4.1). If one does not impose that empirical discipline on MPCs, model comparisons are muddied here because the rational model produces counterfactually low average MPCs while our model produces more realistic and higher ones. So in an undisciplined comparison our model can deliver greater stabilization effectiveness simply because transfers trigger more spending relative to a model with counterfactually low MPCs.

[9] Consistent with our model's prediction that overconfidence limits the effectiveness of targeted transfers, D'Acunto et al. (2023a) find empirically that cognitive constraints can limit the effectiveness of policies targeting household consumption and financial condition.

key micro features of cognitive skills heterogeneity and facilitates quantitative study of macro dynamics, the wealth distribution, and policy design and effectiveness.

A second strand focuses on differences between perceived vs. actual idiosyncratic risk. So far, this literature has focused on reduced-form beliefs and a subset of important macro applications. Balleer et al. (2022) show that working-age individuals in the U.S. are "vastly over-optimistic about their own labor market prospects" (p. 1). Mueller et al. (2021) find optimistic bias about job-finding rates, especially for the long-term unemployed, and little evidence for downward revision of these beliefs when remaining unemployed. Caplin et al. (2023) document that the subjective earnings risk is substantially lower than earnings risk estimated from administrative data. Wang (2023) shows how calibrating a standard incomplete-markets model to consumers' perceived rather than actual income risk is better able to account for observed wealth inequality. Rozsypal and Schlafmann (2023) find that lower-income households tend to underestimate their future income growth. Our contributions are uncovering the role of cognitive skills heterogeneity in shaping biased perceptions about risk and future financial situations, and building a general equilibrium model that can jointly fit key features of micro and macro data and quantitatively evaluate and guide policy.

Third, we contribute to the development of macro models seeking to use insights from behavioral economics to improve predictive and prescriptive power. Most work in this vein focuses on a representative behavioral agent (e.g., Woodford (2013), Gabaix (2014), Woodford (2019), Gabaix (2020), Bordalo et al. (2020), and Lian (forthcoming)). Behavioral HANK models tend to allow for heterogeneity only in the budget constraint, with a homogeneous behavioral or information friction about an aggregate variable only (e.g., Farhi and Werning (2019), Auclert et al. (2020), Angeletos and Huo (2021), Laibson et al. (2021), Pfäuti and Seyrich (2022)). Pfäuti and Seyrich (2022) do study a case of heterogeneous behavioral biases, but focus on expectations about aggregate variables in that case. Guerreiro (2023) allows for heterogeneous attention at the micro level, but focuses on a case where households hold rational expectations about their idiosyncratic shocks. Ilut and Valchev (2023) develop a model of imperfect reasoning and introduce this into an Aiyagari (1994) economy. Their households do not know their optimal policy function and need to estimate it based on costly (and noisy) deliberation signals. In contrast to our framework, households are ex-ante identical. As in standard rational HA(NK) models, HTM status in Ilut and Valchev (2023) is therefore driven by adverse idiosyncratic productivity shocks, i.e., due to bad luck. Once households become hand-to-mouth, they are then likely to remain persistently HtM because they hold excessively high beliefs about their optimal consumption that induce them to dissave and remain at the borrowing constraint. In contrast, HtM households in our setup are likely to differ systematically from households away from the borrowing constraint, consistent with what we find in the data. Additionally, our model features nominal rigidities and allows for two assets. We also take a step beyond the crucial one of matching key empirical moments by demonstrating use cases for our model: analyzing normative and positive implications for fiscal policy.

A fourth strand considers (permanent) heterogeneity in reduced-form or presumed-classical preferences. Aguiar et al. (2021) find that allowing for heterogeneity in patience and the elasticity of intertemporal substitution helps match several empirical facts about the behavior of HtM households. They suggest that behavioral factors might provide a potential micro-foundation for their modeling choices. Krueger et al. (2016) and Auclert et al. (2020) introduce permanent heterogeneity in patience and—in the case of Auclert et al. (2020)—in average skills to better match wealth inequality data. Kekre and Lenel (2022) show that heterogeneity in risk aversion can help account for observed heterogeneity in portfolio choice. Kaplan and Violante (2022) show that heterogeneity in risk aversion can produce similar results to heterogeneity in discount factors in terms of HtM shares and MPCs. They also show, however, that allowing for heterogeneity in risk aversion or in discount factors does not solve the standard HANK's "missing middle problem" of producing a wealth distribution that is too polarized. We show that allowing for heterogeneity in overconfidence, in contrast, fills in the missing middle. Furthermore, our micro data does not favor patience or risk aversion alone as an empirically likely key margin of heterogeneity; e.g., their correlations with HtM status are relatively weak compared to cognitive skills and overconfidence, both qualitatively and quantitatively.

Overall, one of our key contributions is accounting for observed *systematic* differences between financially constrained and unconstrained consumers in a general equilibrium model. This contrasts sharply both with models assuming rational expectations ("RE") and with behavioral models where the only potential deviation from RE regards some aggregate variable. In those classes of models, households become borrowing constrained because they are unlucky, i.e., hit by adverse productivity shocks, and HtM tends to be a relatively transitory state. In our model, households are borrowing constrained because they overestimate their own abilities, leading to a systematic relationship between cognitive skills, overconfidence and persistent HtM status. Our model thus aligns better with empirical findings that liquidity is the main predictor of MPCs (e.g., Fagereng et al. (2021), Jappelli and Pistaferri (2014)). And it turns out that accounting for the underlying reason why some households systematically hold few liquid assets matters greatly for policy.

Outline. We discuss our data and our empirical findings in Section 2. In Section 3, we show how we introduce cognitive skills and overconfidence in HANK, and we present the stationary equilibrium effects of heterogeneity in skills and beliefs in Section 4. The fiscal policy implications are presented in 5 and Section 6 concludes.

2 Micro Data and Empirical Results

In this section, we document several new facts regarding consumers' cognitive skills, beliefs about these skills and future financial situations, and how they relate to other forms of persistent heterogeneity and to six measures of hand-to-mouth status. We later use these facts to help discipline and test our model.

2.1 Data

Our micro data source is the American Life Panel, a long-running online panel that goes to great lengths to obtain a nationally representative sample of U.S. adults.

We measure cognitive skills and overconfidence about cognitive skills using data from the modules in Stango and Zinman (2022, 2023), henceforth SZ, which elicited behavioral biases and cognitive abilities, together with questions about household financial condition (that we use here to construct some of our measures of HtM status), from the same 845 panelists in two survey rounds administered in 2014 and 2017. The SZ modules sample only working-age adults (aged 18-60 in 2014), which maps well into our model's focus on labor-market productivity. We bring in additional variables—regarding standard measures of HtM status not covered in the SZ modules, and standard measures of consumer sentiment that we use to measure subjective financial condition and expectations thereof—using various other ALP surveys administered from 2010 through 2022. We start by detailing our key variable definitions and prevalences, including comparisons to other work where applicable. We then describe the key micro empirical regularities that shape and discipline our model.

Cognitive skills. We measure cognitive skills for SZ panelists with standard tests for fluid intelligence (McArdle et al. (2007)), numeracy (Banks and Oldfield (2007)), cognitive control/executive function (MacLeod (1991), Miyake and Friedman (2012)), and crystalized intelligence in the form of financial literacy (Lusardi and Mitchell (2014)).[10] We then extract a single common factor (a.k.a. "g" or generalized intelligence), to use as a summary statistic for cognitive skills, as is customary given that various cognitive skills measures are strongly related, both conceptually and empirically (Jensen, 1998; Stango and Zinman, 2023).[11]

Overconfidence. To measure overconfidence, we use the responses to the question the SZ panelists are asked: "... what you think about your intelligence as it would be measured by a standard test. How do you think your performance would rank, relative to all of the other ALP members who have taken the test?", elicited as an integer percentile. Later in that survey they take a standard 15-question "number series" test of fluid intelligence (McArdle et al. (2007)).[12] Respondents are overconfident on average, with 70 percent providing a better-than-average percentile.[13]

[10]For details on test questions, please see the Data Appendix to Stango and Zinman (2023).

[11]Results are very similar, qualitatively and quantitatively, if we use the first principal component of cognitive skills instead of the first common factor.

[12]Number series scores correlate strongly with those from other fluid intelligence tests like IQ and Raven's.

[13]The SZ data provides a second measure of (over)confidence about cognitive skills, regarding absolute performance on the numeracy test, that is strongly correlated with our measure of overconfidence in relative performance (Stango and Zinman (2023), Chapman et al. (2023)). We focus on the relative overconfidence measure because it is more powerful, both statistically (it is more granular) and conceptually (fluid intelligence is linked more strongly

We are most interested in heterogeneity in overconfidence and measure it in two ways. One is the degree of overconfidence, defined as the self-assessed rank minus the actual rank so that a higher value of this "oc percentile rank" indicates more overconfidence. The second maps into a key model input: the population share of households exhibiting persistent overconfidence. To estimate this input we flag the 38 percent of respondents who are above-median rank in both 2014 and 2017 as "oc in both rounds" (the standard error on this prevalence estimate is 4pp).[14]

We are not aware of any other quantitative estimate of the share of consumers who are persistently overconfident about their ability, or some closely related object, in a plausibly representative national sample of the working-age population. Huffman et al. (2022) estimate that 45 to 48 percent of managers are over-confident about their performance in a repeated high-stakes workplace tournament held by a single employer. Moschini et al. (2023) find widespread over-optimism about college completion among 18 year-olds in the 1997 NLSY. Various theories explain how overconfidence can persist even in the presence of feedback (e.g., Heidhues et al. (2018) or Zimmermann (2020)).

Subjective financial condition forecasts and realizations. A key innovation of our paper is that we link overconfidence about cognitive skills to consumers' forecasts of their future financial situation. The ALP elicits such forecasts, and subsequent realizations, in many of its survey modules, allowing us to build a panel of 17,266 forecast-realization pairs, provided by 3,401 ALP panelists (including many SZ panelists, as detailed below), across fourteen surveys administered in January and July from July 2010 to January 2016.

The ALP elicits forecast with a question that has long been used, by the Michigan Survey of Consumers and many other national household surveys across the world, to help measure consumer sentiment (e.g., Souleles (2004)): "... do you think that a year from now you will be better off financially, or worse off, or just about the same as now?". These forecasts are highly correlated with expected income growth in the relatively small number of ALP surveys that also elicit an income forecast (Appendix Table A1). We measure realizations a year later with "We are interested in how people are getting along financially these days. Would you say that you are better off or worse off financially than you were a year ago?". Both forecasts and forecast errors tilt strongly optimistic in the aggregate, regardless of the time period (Appendix Table A2).[15] Forecast errors

to productivity than numeracy is).

[14]Data limitations preclude us from estimating prevalence more precisely, by directly comparing each respondent's forecasted to actual percentile, because the forecast's integer percentile support is much more granular than the 15-question test realization's support.

[15]Appendix Table A2 shows that forecasts are more than twice as likely to predict improvement (27 to 30 percent of observations) as deterioration (10 to 14 percent of observations). Forecast errors are roughly three times more likely to be in an optimistic than pessimistic direction; to see this, focus on the "same" realization column to allow for the possibility of forecast errors in either direction, and note that an estimated 13 to 18 percent of the sample forecasted better and ended up the same, while only 4 to 7 percent forecasted worse and ended up the same.

are persistent,[16] and there is only modest evidence of learning over longer periods of time.[17] Nor is there evidence of substantial overcorrection.[18]

Being especially interested in persistent heterogeneity across consumers, we construct three household-level measures of persistent optimism about financial situations. The first one is a dummy variable that equals one if consumer ever makes consecutive optimistic forecast errors. The second is a dummy variable equaling 1 when the proportion of optimistic forecast errors is larger than 0.5. The third is the proportion of optimistic forecast errors. Limiting the sample to forecast-realization pairs with realizations of "about the same" to allow for potential forecast errors in either direction, and to panelists with at least two such pairs, we estimate that 12 to 18 percent of the sample are persistently optimistic in the SZ overlap sample of 409 (Table A5 Columns 5 and 6). The SZ sample is key for our subsequent analysis because we have the requisite measures of overconfidence about cognitive skills only for those panelists. We obtain similar estimates of persistent optimism in the broader ALP sample of panelists with multiple pairs of forecast-"about the same" realizations.

Hand-to-Mouth status. To assess whether someone is (persistently) HtM, we use six different measures of financial constraints. Some of them have been used in previous work, others are new. Two of the six measures are from the two SZ modules. The other four we pull in from other survey modules completed by SZ respondents, so that we can link their HtM status to their cognitive skills and overconfidence thereon.

We start by detailing the two HtM measures from the SZ modules. For both of these, we create indicators for whether someone exhibits the symptom of HtM status in both 2014 and 2017. The first measure indicates *severe financial distress*, defined as reporting that any of the four events happened in the previous 12 months: forced move, late payments, hunger, or foregone medical care. An estimated 28-31 percent of our sample exhibits this indicator in both 2014 and 2017 (for standard errors on these and other estimates of HtM prevalence see Table 1 Columns (7) and (8)). Our second measure classifies a household to be HtM if its liquid net worth is less than half of total monthly household income. About 40-47 percent of our sample exhibits this indicator in both 2014 and 2017. Kaplan and Violante (2022) obtain a similar estimate, of 41 percent, from the 2019 Survey of Consumer Finances.

The third measure of HtM status is indicating strong agreement with the statement: 'I live from paycheck to paycheck' in a 2012 survey. An estimated 56-59 percent of our sample does so. Our

[16] Appendix Table A3 shows that about 75 percent of consecutive forecast errors are the same (both optimistic, both realistic, or both pessimistic), and that 45 percent of panelists who make an optimistic forecast error in the previous period make the same error in the next period.

[17] Comparing the first to last forecast-realization pair we observe for panelists with multiple pairs, Appendix Table A4 shows that the accuracy rate increases from 55 to 62 percent and the optimistic slant decreases from $16/21 = 77$ percent to $13/18 = 72$ percent.

[18] Appendix Table A3 shows that optimists are about 32 times more likely to get better-calibrated than to over-correct with a pessimistic forecast error.

fourth measure is closely related and draws on two questions asked in nine COVID-era modules administered May 2020-July 2022. The mean proportion of these modules in which a panelist exhibits paycheck-to-paycheck behavior is about 40-44 percent.[19] Our fifth measure indicates whether someone lacks precautionary savings, defined as reporting not having emergency or rainy day funds set aside to cover 3-months' of expenses. An estimated 63-69 percent of respondents who completed both surveys where this question was asked indicate this in 2012 and 2018. Our sixth measure is based on whether the panelist indicates having difficulty covering expenses, measured as the proportion of 3 surveys from 2011, 2012, and 2018 where they do not express the highest confidence or certainty that they could cover an unexpected $2,000 need arising in the next month. The mean proportion across panelists is about 51-54 percent, as compared to Sergeyev et al. (2023) who estimate that 54 percent of U.S. households would have difficulty covering an unexpected $2,000 dollar emergency expense in 2022.

Overall, our estimates of HtM prevalence square well with those from prior work. They also suggest that we have measures of financial constraints of varying severity, which will be useful for exploring the robustness of our results below.

2.2 Key Correlations

We now turn to using the above measures of our variables to estimate the key empirical relationships that shape and discipline our model.

2.2.1 Empirical strategy

In estimating empirical relationships between variables, we focus on pairwise correlations, for two reasons. One is empirical: pairwise correlations are easier to interpret when all of the variables of interest are correlated with each other; conversely, multi-variate estimates are likely subject to confounds from over-controlling and multi-collinearity. The other is conceptual: for modeling purposes, we are interested in identifying a proxy for persistent and relatively fundamental consumer heterogeneity (like overconfidence about cognitive skills) that can reproduce key empirical patterns in the aggregate (like patterns of forecast errors and financial constraints). The proxy can be useful, for modeling purposes, whether or not it has a causal relationship with the other variables of interest. We show both unweighted and sampling probability-weighted estimates, following Solon et al. (2015). We address measurement error in cognitive skills, overconfidence, and other potential sources of fundamental and persistent heterogeneity in decision making by using SZ's repeated measurements as instruments for each other where advisable, following Gillen et al.

[19]For each panelist-survey we define an indicator that =1 if panelists respond 'Very difficult' or 'Somewhat difficult' to 'In the past month, how difficult has it been for you to cover your expenses and pay all your bills?' or if on the followup question 'Suppose now you have an emergency expense that costs $400. Based on your current financial situation, how would you pay this expense?' they report one or more expensive options: credit card revolving, small-dollar credit, or that they wouldn't be able to pay for it. For each panelist we then take the ratio of the count of indicators to the count of completed surveys, across the nine modules.

Table 1: Pairwise correlations between persistent HtM measures and cognitive skills, and persistent overconfidence about skills

	CS rank: cf		1=Oc both rounds		Oc pctile rank		Row variable, unw.	Row variable, w.
	Unw.	Weighted	Unw.	Weighted	Unw.	Weighted	Pop. share	Pop. share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Severe financial distress	-0.335	-0.287	0.176	0.273	0.194	0.180	0.277	0.305
s.e.	0.040	0.073	0.059	0.119	0.039	0.078	0.016	0.035
N	841	841	813	813	813	813		
Low net worth	-0.397	-0.368	0.250	0.198	0.226	0.086	0.397	0.468
s.e.	0.038	0.061	0.057	0.097	0.041	0.073	0.018	0.032
N	788	788	760	760	760	760		
paycheck-to-paycheck c. 2012	-0.292	-0.431	0.151	0.023	0.154	0.155	0.588	0.561
s.e.	0.065	0.080	0.099	0.181	0.074	0.099	0.031	0.056
N	263	263	255	255	255	255		
paycheck-to-paycheck, COVID	-0.383	-0.340	0.224	0.220	0.301	0.290	0.404	0.440
s.e.	0.020	0.020	0.053	0.085	0.049	0.077	0.018	0.028
N	2108	2108	516	516	516	516		
No prec. savings in 2012 & 2018	-0.300	-0.339	0.112	0.104	0.181	0.205	0.634	0.691
s.e.	0.070	0.093	0.101	0.133	0.071	0.086	0.030	0.037
N	272	272	262	262	262	262		
Difficult covering \$2k expense	-0.398	-0.446	0.230	0.314	0.222	0.281	0.513	0.543
s.e.	0.041	0.048	0.065	0.078	0.050	0.058	0.021	0.026
N	499	499	485	485	485	485		

Note: CS = cognitive skills, measured as the common factor of four standard tests. Overconfidence oc: relative performance in a cognitive skills test (see Section 2 for details). Weighted estimates use the sampling probability for the last SZ module. In Columns (5) and (6), we use Obviously Related Instrumental Variables to account for measurement error by having the two measurements of oc rank (taken in 2014 and 2017) instrument for each other (Gillen et al. (2019), Stango and Zinman (2023)). We do not take the same approach to the oc indicator in Columns (3) and (4), because measurement error-IV does not work well on misclassification error. Fully non-IV correlations estimated using tetrachoric or Pearson. Details on the exact measurement of the different HtM measures are in the main text. Population shares for the non-indicator variables estimated by taking the mean of the estimated population shares for each survey used in creating that variable.

(2019) and Stango and Zinman (2023).[20]

2.2.2 Cognitive skills and HtM status

As noted at the outset, cognitive skills heterogeneity has been linked to some variables of macroeconomic interest in prior work but not explicitly to HtM status and to its persistence within-household over time.[21] Columns (1) and (2) in Table 1 take steps towards filling that gap. We estimate unweighted and sampling-probability-weighted between our cognitive skills summary measure and each of our six HtM measures, finding a negative sign on all of the 12 point estimates. All of them are larger than |0.20| and most are highly statistically significant with t-stats of |4| or more.

2.2.3 Overconfidence, forecasting, and HtM status

Given that cognitive skills heterogeneity alone is unlikely to help fit the macro data (as we show formally in Section 4), we now consider overconfidence about cognitive skills as a potential mechanism for the observed relationship between cognitive skills and persistent HtM. Indeed, overconfidence in relative performance is the behavioral bias most strongly correlated with cognitive skills out of the 17 biases measured in the SZ data (Stango and Zinman (2023)). Overconfidence could be a key link between cognitive skills and consumer behavior that has been overlooked so far.

Table 2: Optimistic forecast errors are more prevalent among the overconfident

$\frac{\text{(Optimist share} \mid \text{overconfident)}}{\text{(Optimist share} \mid \text{not oc)}}$	Optimism measure	
	1 = (Consec. Opt. FEs)	1 = (Prop. Opt. FEs $\geq$ 0.5)
Unweighted	1.51	1.77
Weighted	1.17	1.63

Note: Sample is the 409 Stango-Zinman panelists who also provide the requisite data, in other ALP modules, on financial condition forecasts and realizations. Overconfidence re: relative peformance in a cognitive skills test (see Section 2 for details). Forecast errors re: household financial condition (see Table A2 for details). Weighted estimates use the mean of each panelist's: (sample probably weight from the last Stango-Zinman module, mean sampling weight across the survey(s) with the realization component of the forecast error(s) used here).

Table 2 links overconfidence about cognitive skills to over-optimism about one's own future financial situation. We see that persistent optimism about one's own future financial condition—as measured by our two indicators—is about 1.5 times more prevalent among the persistently overconfident households than in the rest of the population.[22] In the model calibration, we will

[20]Measurement error IV is advisable for smooth measures but not for discrete ones—the latter is subject to misclassification error that is non-classical.

[21]Recall that 5 of our 6 HtM measures explicitly capture persistence. Because HtM status is so persistent, results on HtM snapshots are similar and not reported below.

[22]Table A5 in the Appendix correlates each of our three measures of persistently optimistic forecast errors with our two measures of persistent overconfidence about cognitive skills, estimating both unweighted and sampling

use this ratio of relative over-optimism to discipline overconfidence.

Table 2 suggests that the strong negative relationship between cognitive skills and HtM status in columns 1 and 2 in Table 1 may be due at least in part to overconfidence. Columns 3-6 in Table 1 provide empirical support for that conjecture. Here we estimate 24 correlations: (6 HtM measures × 2 overconfidence measures × weighted or unweighted). All of the 24 point estimates are positively signed, and eighteen have t-stats strictly greater than two.

2.3 Other sources of fundamental heterogeneity?

Other papers have put forth more-classical sources of relatively fundamental heterogeneity as candidates for macro modeling; see e.g., Auclert et al. (2020), Aguiar et al. (2021), Kaplan and Violante (2022), and Andreou et al. (2023) on patience, and Kaplan and Violante (2022) and Kekre and Lenel (2022) on risk aversion. But we find that the micro data favors focusing on cognitive skills and overconfidence over patience or risk aversion. Stango and Zinman (2023)'s findings point to cognitive skills heterogeneity as the most likely source or summary statistic for heterogeneity in various behavioral biases, and moreover show that overconfidence in relative performance is the bias that has the strongest correlation with cognitive skills. Here we look directly at relationships between our other key micro variables for macro modeling on the one hand, and patience or risk aversion on the other. We do not find evidence of a robust relationship between those classical decision inputs and persistent over-optimism about financial condition, subject to the caveat that any nulls are imprecisely estimated (Appendix Table A8). Turning to HtM status, although we do find some evidence of potentially meaningful correlations with patience or risk aversion, overall the relationships are less robustly strong across our six HtM measures than they are with cognitive skills or overconfidence, both statistically and quantitatively, and patience has a surprising positive correlation with living paycheck-to-paycheck (Appendix Table A9). Nor is patience a good proxy for overconfidence (Appendix Table A10). Risk aversion might be, but the two different measures of presumed-classical risk aversion in the SZ data have opposite-signed correlations with overconfidence, despite being positively correlated >0.2 with each other.

2.4 Summary of results from micro data

To summarize, we find that persistent HtM status decreases strongly with cognitive skills and increases with overconfidence thereon, and that overconfident consumers tend to be persistently too optimistic about their future financial situation. Together with prevalent overconfidence, and the strong negative correlation between cognitive skills and overconfidence found in prior work, these findings suggest that accounting for consumer heterogeneity in cogntive skills and/or

probability-weighted correlations. All 12 of these correlations are positive, which has almost zero probability of happening by chance. The magnitude of the estimated correlations is modest—all but one point estimate fall in the 0.08 to 0.22 range—but this strikes us as unsurprising given the measures' coarseness. The unweighted estimates are more precise, with t-stats ranging from 1.7 to 2.3. The weighted estimates have larger standard errors but similar point estimates for four of the six pairs of results.

overconfidence could be important for macroeconomic fluctuations. We next develop a model that can explain these findings and analyze how they matter for fiscal policy.

3 Model

In this section, we present a HANK model with permanent heterogeneity in cognitive skills and overconfidence about these skills. The model features incomplete markets in the spirit of Bewley (1986), Huggett (1993), and Aiyagari (1994), and nominal rigidities in the form of sticky wages. Time is discrete and denoted by $t = 1, 2, ...$ We first focus on the case in which households can only save in one asset; a liquid bond issued by the government. Later on, we introduce a second asset in the form of illiquid productive capital.

Households. There is a unit mass of households that are subject to idiosyncratic risk, incomplete markets, and borrowing constraints. We allow for permanent heterogeneity in households' cognitive skills (or productivity) and overconfidence about these cognitive skills (specifically about idiosyncratic productivity), consistent with our empirical measure of overconfidence in Section 2. An individual household's skills of permanent type g in period t are denoted by $\bar{e}_g e_t$, where $\bar{e}_g$ captures permanent differences across groups in average skill levels, and e_t captures idiosyncratic fluctuations in skills. The stochastic component e_t follows a Markov process with time-invariant transition matrix $\mathcal{P}$. The process for e_t is the same for all households and the mass of households in state e is always equal to the probability of being in state e in the stationary equilibrium, $p(e)$.

The problem of an individual household of type g in idiosyncratic state e_t, with beginning-of-period liquid asset holdings b_{t-1}, is given by:

$$V_{g,t}\left(b_{t-1}, e_t\right) = \max_{c_t, b_t} \left\{ \frac{c_t^{1-\gamma}}{1-\gamma} - \frac{n_t^{1+\varphi}}{1+\varphi} + \beta \tilde{\mathbb{E}}_{g,t} V_{g,t+1}\left(b_t, e_{t+1}\right) \right\}$$

subject to

$$c_t + \frac{b_t}{1+r_t} = b_{t-1} + (1-\tau_t) w_t \bar{e}_g e_t n_t \tag{1}$$

$$b_t \geq -\underline{b}, \tag{2}$$

where c_t denotes consumption, n_t hours worked, r_t the net real interest rate, w_t the real wage, τ_t denotes the income tax rate, and V the value function. The parameters γ, φ, and β denote the relative risk aversion, the inverse Frisch elasticity of labor supply, and the time discount factor, respectively. These parameters as well as the exogenous borrowing limit $\underline{b}$ are the same for all households and are time-invariant.

The expectations operator $\tilde{\mathbb{E}}_{g,t}$ depends on g, which not only captures permanent heterogeneity in cognitive skills but also in overconfidence. Overconfidence affects the perceived future cognitive skills, as we discuss next.

Cognitive skills and overconfidence. We allow for permanent heterogeneity in cognitive skills and overconfidence about these cognitive skills. Heterogeneity in cognitive skill levels is modelled as different average productivities $\bar{e}_g$, given the strong (negative) correlation between cognitive skills and income in the data, as we discuss in more detail in the calibration section later (see also Appendix Table A6).

All households observe their current cognitive skills $\bar{e}_g e_t$ but overconfident households have biased beliefs about the transition probabilities $p(e_{t+1}|e_t)$. Specifically, overconfident households assign too much probability to reaching (or staying in) relatively high-skill states, and too little probability to reaching (or staying in) relatively low-skill states. This makes overconfident households too optimistic about their expected future cognitive skills, relative to a rational household with the same cognitive skills and idiosyncratic risk.

Let $p_{ij} \equiv p(e_{t+1} = e_j | e_t = e_i)$ denote the probability that a household with current skill level $e_i \in \{e_1, e_2, ..., e_J\}$ reaches skill level $e_j \in \{e_1, e_2, ..., e_J\}$ in the following period, and assume that the skill levels are ordered such that $e_1 < e_2 < ... < e_J$. To capture overconfidence with only one additional parameter independent of the number of skill states, we assume that an overconfident household's *perceived* transition probabilities $\tilde{p}_{ij}$ are given by

$$\tilde{p}_{ij} \equiv \begin{cases} \alpha p_{ij}, & \text{if } i < j \\ \frac{1}{\alpha} p_{ij}, & \text{if } i > j \\ 1 - \sum_{j \neq i} \tilde{p}_{ij}, & \text{if } i = j, \end{cases} \tag{3}$$

where the parameter $\alpha \geq 1$ captures overconfidence. If $\alpha > 1$, the household assigns too much weight to reaching a better state (this is the case $i < j$) and too little weight to reaching a worse state ($i > j$). The perceived probability of staying in the same state ($i = j$) ensures that the probabilities sum to 1.[23] We discuss an alternative modelling approach in Section 4.3, where the degree of overconfidence depends on the distance of the states. Note that the rational expectations case is captured by setting $\alpha = 1$ and thus nested in our setup.[24]

An immediate implication is that overconfident households will more often be overly optimistic about their financial situation (specifically income, in the model) compared to rational households, consistent with the empirical findings reported in Section 2 (Tables A5 and 2). In Section 3.1 below we target our empirical estimate of the relative share of optimists among overconfident and rational households from Table 2 to calibrate α.

[23] We further restrict α such that all perceived transition probabilities lie between 0 and 1. Given a standard calibration for the income process, this restriction is never binding.

[24] Modelling overconfidence as in (3) is similar to the way Caballero and Simsek (2020) model optimism about an aggregate state with two possible realizations. In contrast to them, we focus on idiosyncratic states and allow for an arbitrary number of realizations. McClung and Nighswander (2021) introduce belief heterogeneity about idiosyncratic employment transition probabilities into a life-cycle model, but consider only two possible states.

Unions. We follow the recent HANK literature and assume that hours worked n_t are determined by union labor demand and that wages are sticky whereas prices are flexible (most closely to our setup, see Auclert et al. (2023), which is based on Erceg et al. (2000)).[25] Each worker provides $n_{k,t}$ hours of work to a continuum of unions indexed by $k \in [0, 1]$. Each union aggregates efficient units of work into a union-specific task

$$N_{k,t} = \int \bar{e}_i e_{i,t} n_{i,k,t} di,$$

where i here denotes the individual household and thus, indicates both its permanent type as well as its current idiosyncratic state.

A competitive labor packer then packages these tasks into aggregate employment services according to the CES technology

$$N_t = \left(\int_k N_{k,t}^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \tag{4}$$

and sells these services to firms at price w_t.

We model wage stickiness by imposing a quadratic utility cost $\frac{\psi}{2} \int_k \left(\frac{W_{k,t}}{W_{k,t-1}} - 1 \right)^2 dk$ that shows up in the households utility function. A union sets a common nominal wage $W_{k,t}$ per efficient unit for each of its members.

In doing so, the union trades-off the marginal disutility of working given average hours against the marginal utility of consumption given average consumption. The union then calls upon its members to supply hours. We assume the union ensures that all households supply the same amount of hours, $n_{i,t} = N_t$ for all households i at all times t.

Firms. A representative firm operates an aggregate production function which is linear in labor input N_t

$$Y_t = N_t, \tag{5}$$

to produce total output Y_t. Prices are fully flexible such that the real wage per efficient hour is constant

$$w_t = 1. \tag{6}$$

Profits are zero. Since the nominal wage is given by $W_t \equiv w_t P_t = P_t$, we have

$$1 + \pi_t = 1 + \pi_t^w, \tag{7}$$

where $\pi_t \equiv \frac{P_t}{P_{t-1}} - 1$ denotes goods price inflation, and $\pi_t^w \equiv \frac{W_t}{W_{t-1}} - 1$ wage inflation.

[25] Auclert et al. (2021) and Broer et al. (2020) argue in favor of using sticky wages rather than sticky prices in HANK models.

Fiscal policy. We abstract from government spending and assume that the fiscal authority sets total taxes minus transfers, T_t, following a simple debt feedback rule

$$T_t - \bar{T} = \vartheta \frac{B_t - \bar{B}}{\bar{Y}}, \tag{8}$$

where $\bar{T}$, $\bar{B}$ and $\bar{Y}$ denote the stationary equilibrium values of taxes, government debt and output, respectively. Furthermore, the government budget constraint is given by

$$B_t + T_t = R_t B_{t-1}. \tag{9}$$

Monetary policy. The monetary authority directly controls the real rate r_t and we assume that they keep it constant at its steady state value r at all times. Note, that this assumptions does neither matter for the stationary equilibrium analysis nor the optimal government debt exercise that we discuss in Section 4 and 5. The assumption only matters when we consider aggregate shocks, as we do in Section 5.2.2, when we examine the role of overconfidence for the effectiveness of temporarily increasing fiscal transfers.

Equilibrium. Absent aggregate shocks, and given an initial price level P_{-1}, initial nominal wage W_{-1}, initial government debt B_{-1}, and an initial distribution of agents $\Psi_{g,0}(b_{-1}, e_0)$ in each fixed group g, a general equilibrium is a path for prices $\{P_t, W_t, \pi_t, \pi_t^w, r_t, i_t\}$, aggregates $\{Y_t, C_t, N_t, B_t, T_t\}$, individual allocation rules $\{c_{g,t}(b_{t-1}, e_t), b_{g,t}(b_{t-1}, e_t)\}$ and joint distributions of agents $\Psi_{g,t}(b_{t-1}, e_t)$ such that households optimize (given their beliefs), all firms optimize, unions optimize, monetary and fiscal policy follow their rules, and the goods and bond markets clear:

$$\sum_{g,e} \mu_g p(e) \int c_t \Psi_{g,t}(b_{t-1}, e_t) = Y_t \tag{10}$$

$$\sum_{g,e} \mu_g p(e) \int b_t \Psi_{g,t}(b_{t-1}, e_t) = B_t, \tag{11}$$

where μ_g denotes the mass of agents of type g.

3.1 Calibration

One period in the model corresponds to a quarter. We calibrate the standard parameters to values often used in the literature. For the idiosyncratic skill process, we follow McKay et al. (2016) and set the autocorrelation of e_t to $\rho_e = 0.966$ and the variance to $\sigma_e^2 = 0.016$. We then discretize this process into an eleven-states Markov chain using the Rouwenhorst (1995) method. We set the discount factor, β, to match a steady state real interest rate of 4% (annualized). Risk aversion is set to $\gamma = 2$, the inverse Frisch elasticity to $\varphi = 2$, and the borrowing limit to $\underline{b} = 0$ (as, e.g., in McKay et al. (2016)). We set the average wealth to average annual income ratio to 4.1 as in

Kaplan and Violante (2022).

Table 3: Persistent overconfidence: prevalence and relationship to income

	Overconfident in both survey rounds?			
	Yes Unweighted	No Unweighted	Yes Weighted	No Weighted
Population share	0.34		0.38	
	(0.02)		(0.04)	
Mean Income	51,182$	79,765$	42,035$	77,145$
N	817	817	817	817

Note: Standard errors in parentheses. Weighted estimates use the sampling probability for the last SZ module.

Our key innovation is permanent heterogeneity in cognitive skills and in overconfidence. We set the share of overconfident households to 0.38 as estimated in the data (see the upper part in Table 3). Based on prior work showing strong negative correlations between cognitive skills and overconfidence about those skills (see Ehrlinger et al. (2008); Stango and Zinman (2023) or Table A7), for now we assume that all overconfident households are low-skilled and collapse permanent heterogeneity in skills and confidence to two types: overconfident with low skills, and rational with high skills. Using sampling probability weights, we estimate that the average income of overconfident households is about 42,000 USD, whereas it is about 77,000 USD for the households that are not overconfident (see lower part in Table 3). To match this ratio of $0.55 = \frac{42{,}000}{77{,}000}$, we normalize the average productivity of the high-skilled and rational households to $\bar{e}_2 = 1$ and set the average skill level of the low-skilled and overconfident households to $\bar{e}_1 = 0.55$.

Following equation (3), we capture the degree of overconfidence in the overconfident and low-skilled group with one parameter, α. To calibrate α, we target our estimate from Table 2 that overconfident households are 1.5 times as likely to have optimistic one-year forecast errors about their financial situation.[26] This results in $\alpha = 1.9$. Below we consider other parameterizations of heterogeneity in cognitive skills and overconfidence, including one in which all households are rational and differ only in their average skill levels, and one in which some households are under-confident (Sections 4.1 and 4.3).

Table 4 summarizes the calibration of our model with heterogeneity in cognitive skills and overconfidence.

4 Stationary Equilibrium Predictions

We now consider the ability of our model with heterogeneity in cognitive skills and overconfidence to fit various key moments from macro and micro data, as compared to HANK models abstracting

[26] Note, that even though α operates on the perceived transition probabilities of idiosyncratic productivity, households are equally likely to be overly-optimistic about their future productivity as about their financial situation (defined as labor income plus asset income). Thus, modelling overconfidence about skills rather than financial situations as a whole is without loss of generality in our setup.

Table 4: Stationary equilibrium calibration

Parameter	Description	Value
R	Steady state real rate (annualized)	4%
γ	Risk aversion	2
φ	Inverse of Frisch elasticity	2
$\underline{b}$	Borrowing limit	0
$\frac{\bar{B}}{4\bar{Y}}$	Average wealth to average income	4.0
Idiosyncratic risk		
ρ_e	Persistence of idiosyncratic risk	0.966
σ_e^2	Variance of idiosyncratic risk	0.016
Permanent heterogeneity		
μ_g	Mass of households	$\{0.38, 0.62\}$
$\bar{e}_g$	Cognitive skills	$\{0.55, 1\}$
α	Degree of overconfidence	1.9

Note: This table summarizes the calibration of our one-asset model with two groups of permanent heterogeneity: group one has relatively low average skill levels $\bar{e}_1 < \bar{e}_2$ and households in that group are overconfident ($\alpha > 1$), whereas households in group two are relatively high skilled and have rational expectations ($\alpha = 1$).

from either cognitive skills or belief heterogeneity or both.

4.1 Hand-to-Mouth Shares and Average MPCs

We start by considering the effects of permanent heterogeneity in cognitive skills and overconfidence on the share of Hand-to-Mouth (HtM) households and the implied average marginal propensity to consume (MPC) of households.

Table 5 compares predictions across four different models: our baseline model with heterogeneity in cognitive skills and overconfidence (*"HANK: CS + OC"*, in column (1)), a standard HANK model (column (2)) with no heterogeneity in permanent productivity levels ($\bar{e}_g = \{1, 1\}$) and full rationality ($\alpha = 1$), a HANK model with permanent heterogeneity in skill levels but full rationality (*"HANK: CS"*, column (3)), and a HANK model with a group of permanently overconfident households but no skill heterogeneity (*"HANK: OC"*, column (4)).[27] We start by comparing our model to the standard HANK, and then use the other two models to help unpack the differences.

Column (2) reproduces the well-documented finding that a standard one-asset HANK model calibrated to match average wealth produces an average MPC and a HtM share that are both far below consensus estimates (Auclert et al. (2023), Kaplan and Violante (2022)). The reason is that rational households have a high incentive to self-insure themselves against their idiosyncratic risk by accumulating liquid wealth. Thus, with a high enough liquidity supply in the economy, almost no households end up being at the borrowing constraint.

[27] When comparing these four different models, we always recalibrate the discount factor such that all models have the same asset supply and the same steady-state real interest rate. The rest of the calibration is the same for all models.

Table 5: MPCs and shares of HtM households across the models.

	HANK: CS + OC (1)	Standard HANK (2)	HANK: CS (3)	HANK: OC (4)
HtM Share	0.252	0.02	0.020	0.246
Avg. MPC	0.164	0.033	0.032	0.181
HtM rational HHs	0.021	0.020	0.021	0.010
Avg. MPC rat. HHs	0.032	0.033	0.034	0.022
HtM OC HHs	-	-	-	0.632
Avg. MPC OC HHs	-	-	-	0.441
HtM rat. HHs Low-Skilled	-	-	0.020	-
Avg. MPC rat. HHs LS	-	-	0.030	-
HtM OC HHs LS	0.643	-	-	-
Avg. MPC OC HHs LS	0.390	-	-	-

Note: MPCs refer to MPCs out of a \$500 dollar stimulus check. "HANK: CS + OC" is our model in which we allow for skill heterogeneity and overconfidence, "Standard HANK" denotes a standard one-asset model, in which we abstract from heterogeneity in skills and overconfidence, "HANK: CS" denotes the same model, but in which we allow for heterogeneity in skills, and "HANK: OC" denotes a model in which we only allow for overconfidence but not for skill heterogeneity.

In contrast, our model with skill and belief heterogeneity (column (1)) produces an average MPC and a HtM share that are both an order of magnitude larger than in the standard one-asset HANK model. Our predictions align well with consensus estimates, albeit more obviously so for the MPC. For example, Jappelli and Pistaferri (2010) and Havranek and Sokolova (2020) report average MPC estimates in the range of 15-25% over a quarterly time horizon, as compared to our 16.4%. Our predicted share of HtM households, 0.252, is closest to our estimated empirical share based on our most conservative definition of HtM status: those with severe financial distress (Table 1). This is reasonable given our model's definition of HtM is also strict: HtM households are those who are exactly at the borrowing constraint.

Column (3) shows that skill heterogeneity alone does not drive our model's ability to fit the data better. If we introduce only skill heterogeneity but keep all households rational (well-calibrated about their productivity), column (3) shows that the average MPC and the HtM share are very similar to those produced by the standard HANK model. The reason is that a rational household still has a strong incentive to self-insure regardless of its average productivity.

Column (4) shows that our model's allowance for belief heterogeneity drives its improved performance. Specifically, keeping average productivity homogeneous but allowing some households to be overconfident about their future idiosyncratic productivity generates average MPCs and HtM shares that are consistent with the data. The mechanism is that overconfident households overestimate their expected income; i.e., they perceive their income risk to be lower than it actually is. The overconfident households thus accumulate less precautionary savings than rational households facing the same actual income risk. As a result and in line with our empirical findings in Section 2, overconfident households are much more likely to end up being HtM than rational households (64% of overconfident households are HtM, but only 2% of rational households in our model with skill and belief heterogeneity are HtM). This also results in a high average MPC in the group of low-skilled, overconfident households (39% vs. 3% for the rational households) which drives up the aggregate average MPC.

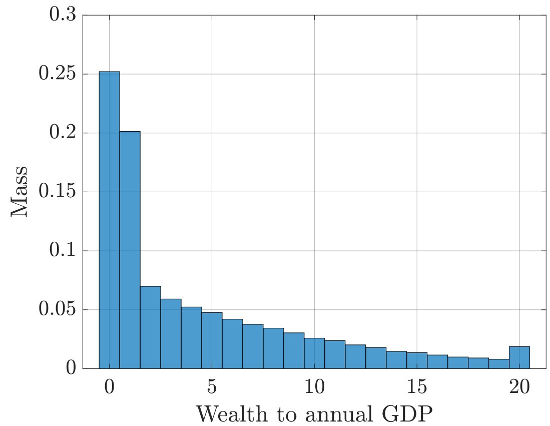
Note that if all agents where overconfident, the model would not deliver these results. If everybody underestimates their income risk in the same way, the price of the asset would decrease until the overconfident households would hold all the assets. As a result, households would be well-insured again. Sufficient *heterogeneity* in overconfidence, and specifically the presence of enough rational households who are relatively willing to save, is required. The high asset demand of the rational households, in general equilibrium, decreases the return on savings and therefore, makes saving less attractive for overconfident households.

4.2 "Missing Middle Problem" and the Top 10% Wealth Share

Standard one-asset HANK models can generate a high average MPC by restricting the amount of wealth in the model (Wolf (2021), Seidl and Seyrich (2021), Kaplan and Violante (2022)). However, on top of not being able to match the amount of average wealth in the economy, producing an empirically realistic MPC by restricting the amount of wealth in the model comes at the cost of an unrealistic wealth distribution, specifically the "Missing Middle Problem" of excessive polarization (Kaplan and Violante, 2022). One way to see this missing middle is that median wealth to mean annual earnings is about a magnitude smaller than in the data. We offer further confirmation of this finding by recalibrating the standard HANK model we use in Table 5 Column (2) to match the average MPC produced by our one-asset model with skill and belief heterogeneity, which requires setting total wealth to income to 0.6 instead of 4.1. This delivers a median wealth to average annual income ratio of 0.2, as compared to about 1.5 in the data (Kaplan and Violante, 2022).

Our one-asset model with heterogeneity in cognitive skills and overconfidence fills in the missing middle: it predicts a median wealth to average income ratio of 1.8, close to albeit slightly above its empirical counterpart of 1.5. Rational households that have experienced several periods of relatively low productivity make up most of the middle of our wealth distribution. Overconfident households tend to be HtM and thus account for most of the bottom, as discussed above. Rational households that have not experienced long spells of bad productivity shocks populate the top of the distribution. Overall, we predict that the top 10% of households hold 38% of total wealth in our economy, which is relatively close to the two-asset model of Kaplan and Violante (2022) and to the empirical target of about 49%. Figure 1 shows the wealth distribution for our model with heterogeneity in cognitive skills and overconfidence.

Figure 1: The wealth distribution



Note: This figure shows the household wealth distribution for our one-asset model with permanent skill and belief heterogeneity.

4.3 Extensions

We now show that our results are robust to: (i) accounting for the empirical finding that 11% of households are persistently *underconfident*, and (ii) an alternative specification of overconfidence in which the degree of overconfidence depends on the household's current idiosyncratic productivity level.

Underconfident households. Our survey data suggests that 11% of households are persistently underconfident in the sense that they underestimate their cognitive skills in both survey rounds. We extend our model to account for this by setting $\alpha_{uc} < 1$ for 11% of households and adding a symmetric target to its calibration: we now not only target overconfident households being 1.5 more likely to be optimistic about their future situations than their rational counterparts, but also underconfident households being 1.5 times more likely to be too pessimistic. The discount factor adjusts again to keep the real interest rate at 4% annually.

Incorporating underconfident households increases the share of HTM slightly from 25.2% to 25.7% and the average MPC from 16.4% to 17.0%. The reason is that underconfident households overestimate their precautionary savings motive compared to rational households. They have an even greater desire than rational households to self-insure their risk. As a result, aggregate savings demand and, thus, the price of savings increases further, which crowds out savings from households close to the borrowing constraint. The existence of underconfident households also pushes up the share of top 10% wealth in the economy to 41% and, thus, closer to its empirical counterpart of 49%.

Overall, extending the model by underconfident households underscores the potential of heterogeneity in beliefs about skills in improving the model's fit to the data. Given the small share of underconfident households in the data, adding them to our model has however only small quantitative effects.

Alternative way of modelling overconfidence. In our baseline specification of overconfidence (equation (3)), the degree of overconfidence is the same for all overconfident households, independent of their current state or skill level. We now allow for dependence of the following form:

$$\tilde{p}_{ij} \equiv \begin{cases} \alpha^{(e_j - e_i)} p_{ij}, \text{ if } i \neq j \\ 1 - \sum_{j \neq i} \tilde{p}_{ij}, \text{ if } i = j. \end{cases} \tag{12}$$

As in our baseline specification, when $\alpha > 1$, the transition probabilities of moving upwards ($e_i < e_j$) are overweighted and the probabilities of moving downward are underweighted. But now these probability distortions are larger for states that are further away from each other.

We again calibrate α to match the empirical finding that overconfident households are 1.5 times as likely to be overly-optimistic about their future earnings compared to rational agents. Here this results in $\alpha = 4.15$. The predicted average MPC is 0.163 and thus largely unchanged from our baseline estimate of 0.164. The predicted HtM share is now about 8 percentage points higher, at 34.4%, and thus closer to the empirical shares of more expansive definitions of HtM (see Table 1).

4.4 Overconfidence in a Two-Asset Model

Rational HANK models often introduce a second, illiquid asset to match the average MPC while simultaneously matching total wealth in the economy (Kaplan et al. (2018), Kaplan and Violante (2022), Auclert et al. (2023)). This approach is meant to capture illiquid assets that are good long-run savings vehicles but ill-suited for self-insurance purposes. But in order to match high average MPCs, two-asset HANK models typically require a liquidity premium—a return difference between liquid and illiquid assets—that is arguably substantially higher than in the data (Kaplan and Violante (2022)).

We now show that the two-asset version of our model can fit the MPC and wealth data with a substantially lower liquidity premium than required by a standard two-asset HANK model.

Model. We add an illiquid asset by enriching the model in two ways. First, households can now save in two assets: one liquid but low-return, and one high-return but illiquid. Second, we model the illiquid asset as productive capital and so the production function now includes capital.

The household's budget constraint now reads:

$$c_t + \frac{b_t}{1 + r_t} + k_t = b_{t-1} + (1 + r_t^k) k_{t-1} + (1 - \tau_t) w_t \bar{e}_g e_t n_t, \tag{13}$$

where k denotes the illiquid asset of the household and r^k is its return. Capital depreciates at rate δ and depreciated capital has to be replaced for maintenance, such that r_t^k is the net return on the illiquid asset. We follow Bayer et al. (forthcoming) and assume that households make their savings and portfolio choices between liquid bonds and illiquid capital in light of a capital market friction: participation in the capital market is random and i.i.d. in the sense that only a fraction λ of households are able to adjust their capital holdings in a given period. Households that do not participate in the capital market ($k_t = k_{t-1}$) still obtain the return on their assets and can adjust their bond holdings. We further assume that both holdings of bonds and holdings of capital have to be non-negative:

$$b_t, k_t \geq 0.$$

A representative firm operates a Cobb-Douglas production function using capital, K, and labor, N, as input factors:

$$Y_t = K_{t-1}^{\chi} N_t^{1-\chi}, \tag{14}$$

where χ denotes the capital share in production.

In addition to the equilibrium conditions in Section 3, now the capital market also needs to clear:

$$\sum_{g,e} \mu_g p(e) \int k_t \Psi_{g,t}\left(k_{t-1}, e_t\right) = K_t. \tag{15}$$

Calibration. We maintain the same values for each of the parameters that also appear in our baseline model (except for the discount factor). Table 6 shows the calibration of the additional parameters and the discount factor. We set the capital share $\chi = 0.318$ which is a standard value in the literature (Bayer et al. (forthcoming)). We then use the quarterly depreciation rate, δ, the probability to participate in the capital market, λ, and the discount factor, β, to jointly target the average wealth-to-annual income ratio of 4.1 and the liquid asset-to-annual income of 0.2 as in Kaplan and Violante (2022), and the average MPC of 0.186 which is well in the range of empirical estimates.

Table 6: Calibration two-asset model

Parameter	Description	Value
χ	Capital share	0.318
δ	Depreciation rate	0.016
λ	Capital market participation rate	0.22
β	Discount factor	0.9935

Note: This table summarizes the new parameters of the two-asset model. All other parameters stay the same as in our baseline model.

Stationary Equilibrium Results. Table 7 shows the influence of overconfident households on the stationary equilibrium (column (1)). Rational households accumulate liquid assets to self-insure before saving in the illiquid asset. Overconfident households remain much more likely to be HtM (70% vs. 8%) because they foresee little value in accumulating a liquid buffer stock and hence prioritize the illiquid asset's higher return if they do save. Compared to a standard two-asset HANK model (Columns (2) and (3)), this dampened demand for liquidity drives the annual liquidity premium down to 2.2%.

Table 7: MPCs and liquidity spread across two-asset models.

	two-asset HANK with overconfidence	rational two-asset HANK	
	(1)	(2)	(3)
		calibrated as (1)	re-calibrated
HtM	0.318	0.12	0.24
Avg. MPC	0.186	0.077	0.16
return gap (annualized)	2.2%	4.0%	8.2%
HtM rat. HHs	0.082	0.06	0.24
Avg. MPC rat. HHs	0.048	0.058	0.16
HtM OC HHs	0.703	-	-
Avg. MPC OC HHs	0.412	-	-

Note: MPCs refer to MPCs out of a stimulus check of $500. "baseline two-asset" denotes our two-asset HANK model with heterogeneity in skills and with overconfidence, "rational two-asset" is the same two-asset HANK model minus heterogeneity in skills and minus overconfidence, and "two-asset recalib." is the latter model recalibrated such that it has an average MPC of 0.16.

Column (2) keeps all the parameters the same as in our behavioral version of the two-asset model in Column (1), except for recalibrating β to target the mean wealth to annual income ratio of 4.1.[28] This produces too few HtM households and thus an average quarterly MPC, 0.077, that is substantially below the consensus range of empirical estimates. Targeting an average MPC at the lower bound of the empirical estimates, e.g. 0.16, produces a more realistic HtM share but already requires a return gap of 8.2%, which is arguably too high compared to empirical estimates (Kaplan and Violante, 2022).[29] Empirical estimates of the return gap are in the ballpark of 5% (see, e.g., Jordà et al. (2019)). When comparing the model-implied return gaps with the empirically estimated one, note, however, that both models abstract from aggregate risk. Accounting for aggregate risk is likely to push up the model-implied risk premia, bringing the risk premium of our model with skill and belief heterogeneity closer to the data and pushing the one from the rational model even further away from it.

[28]This requires quarterly $\beta = 0.992$, as compared to 0.994 in our model.

[29]In targeting the quarterly average MPC of 0.16 we set $\beta = 0.9895$, $\lambda = 0.065$, and $\delta = 0.0125$.

5 Fiscal Policy

Next we show that accounting for overconfident households matters for fiscal policy. We start by analyzing the effects of liquidity provision through public debt and quantifying the optimal public debt level. We then show that the presence of overconfident households increases the efficacy of targeted transfers as it pushes up the MPCs of households. Low wealth standard HANK models, on the other side, predict an even higher efficacy of targeted transfers due to a highly concentrated share of HTM households at the bottom of the income distribution which is at odds with the data.

5.1 Liquidity Provision and the Optimal Public Debt Level

Government-issued debt is an important fiscal policy instrument in HANK models as it increases the supply of liquid assets. Given that markets are incomplete, households self-insure their idiosyncratic income risk by accumulating liquid bonds. Hence, a higher government debt level increase households' abilities to self-insure and smooth consumption. Yet, a higher government debt level increases the amount of distortionary taxes which are required to finance the government's additional interest rate payments.

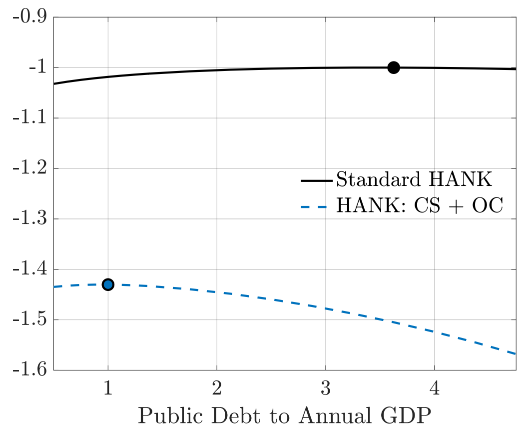
Given this trade-off, we examine the optimal government debt level in our model with heterogeneity in cognitive skills and overconfidence. We use a utilitarian social welfare function that seeks to maximize the average expected discounted lifetime utility of households.[30] The expectations over the individual lifetime utilities in the social welfare function are assumed to be rational, in the spirit of what Benigno and Paciello (2014) call "paternalistic". When varying the supply of government debt, we fix the discount factor β as calibrated in Table 4 and let the interest rate adjust to clear the bond market.

Figure 2 shows that average welfare peaks at a much lower debt level in our model compared to the standard one-asset HANK model: optimal debt is about 100% of annual GDP, compared to about 360% in the standard HANK model.[31] The key driving force is that overconfident households underestimate their income risk and therefore have a dampened response to the liquidity supply increase; indeed, the very households that the government would like to save more are the least responsive ones. This diminishes the average benefit of higher government debt compared to the standard model. Figure 3 illustrates this: the HtM share in panel (a) and especially the bottom-wealth share (panel (b)) are relatively unresponsive to government liquidity provision in our model. Equally strikingly, the share of HtM households plateaus well above zero (e.g., it is about 0.25 at a Debt/GDP ratio of 4), highlighting that even when liquidity is abundant, overconfident households do not tend to save themselves out of being liquidity-constrained. This offers a stark contrast to the standard model, where households at or near the borrowing constraint have the strongest

[30] Such an objective function takes into account aggregate efficiency, risk-sharing, and intertemporal-sharing (Dávila and Schaab, 2022).

[31] For readability, we normalize welfare such that the highest level of welfare in the model with rational expectations is normalized to -1.

Figure 2: Welfare

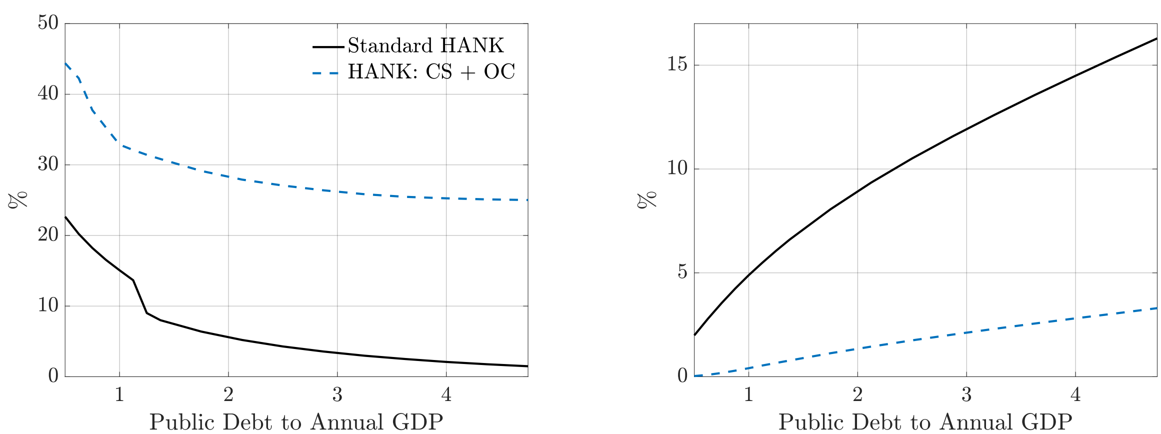


Note: This figure shows the average welfare defined as average expected discounted lifetime utility as a function of government debt and the dots show the welfare-maximizing amount of government debt for our baseline model (blue-dashed lines) and the rational counterpart (black-solid line). The y-axis shows (normalized) average expected lifetime utility, and the x-axis shows public debt to annual GDP, $\frac{B}{4Y}$.

incentive to save in liquid assets and respond strongly as the price of liquidity falls, driving their wealth share up (e.g., to almost 15% at a Debt-to-GDP ratio of 4, compared to about 3% in our model with cognitive skills and overconfidence) and their HtM likelihood down.

Figure 3: The Implications of Higher Government Debt

(a) HtM Share (b) Bottom 50% Wealth Share



Note: This figure shows the share of HtM households in panel (a) and the wealth share of the bottom 50% of households in panel (b) for varying degrees of average government debt to average earnings ratios (horizontal axis). The black-solid lines show the case for the one-asset standard HANK model that abstracts from permanent heterogeneity in cognitive skills and overconfidence, and the blue-dashed lines show the case for our baseline HANK model featuring permanent heterogeneity in cognitive skills and overconfidence.

Overall, Figures 2 and 3 illustrate how making saving more attractive for households is not as powerful once we account for households' overconfidence about their idiosyncratic risk. The optimal government debt level is thus substantially lower when accounting for heterogeneity in overconfidence.[32]

5.1.1 Optimal government debt in the two-asset model

When extending the analysis to the two-asset model, introduced in Section 4.4, we find again that the optimal government debt level is substantially lower in the model with heterogeneity in overconfidence compared to the rational model—about 20% of annual GDP vs. 45% in the model with rational beliefs. In both two-asset models, more government debt crowds out productive capital, and this makes the optimal debt level lower compared to their one-asset versions. Yet still, the overconfident households' relative unresponsiveness to liquidity supply remains quantitatively important when they can substitute into illiquid savings, such that the optimal debt level is lower than in the model in which all households are rational.

5.2 Targeted Transfers

Another important fiscal policy tool in heterogeneous agent models are targeted transfers. Income-targeted transfers are commonly used both as long-run policies to mitigate households income risk and also as short-run stabilization policies. As the effects of targeted transfers crucially depend on the precautionary savings behavior of households and of their marginal propensities to consume, we now use our model to conduct a positive analysis of income-targeted transfers. We start by modeling how the presence of transfers as public insurance affect the stationary equilibrium. We then examine how unanticipated income-targeted transfers affect macro stabilization.

5.2.1 Targeted transfers as public insurance

For now, we model transfers $tr_{i,t}$ to household i at time t in the style of minimum income benefits following Bayer et al. (2023b). Transfers provide additional resources if pre-tax labor income $w_t n_{i,t} e_{i,t}$ falls short of some threshold level:

$$tr_{i,t} = \max\{0, a_1 \bar{y} - a_2 w_t n_{i,t} e_{i,t}\},$$

where $\bar{y}$ is the median income in the stationary equilibrium and $0 \leq a_1, a_2 \leq 1$. Transfers thus decrease in individual income at the withdrawal rate a_2 and no transfers are paid to households whose labor income satisfies $w_t n_{i,t} e_{i,t} \geq \frac{a_1}{a_2} \bar{y}$. We set $a_1 = 0.5$ and $a_2 = 0.8$. For simplicity, we assume that these transfers do not distort labor supply (following Bayer et al. (2023b)). By reducing the net income risk of households, these transfers provide some public insurance.

Total transfer payments of the government are then given by

$$Tr_t = \mathbb{E}_t tr_{it},$$

[32] Figure 2 also shows that welfare in the standard model is less sensitive to deviations from the optimal government debt level than in our model with belief heterogeneity.

where the expectation operator is the cross-sectional average. These transfers are financed via labor-income taxes.

Table 8: Effects of introducing public insurance

	HANK: CS + OC	Standard HANK	Standard HANK, low wealth
	(1)	(2)	(3)
HtM Share	0.252	0.021	0.208
Avg. MPC	0.164	0.032	0.164
Bottom50W	0.029	0.143	0.027
Real rate	4%	4%	4%
HtM Share with PI	0.26	0.055	0.235
Avg. MPC with PI	0.158	0.045	0.215
Bottom50W with PI	0.024	0.119	0.017
Real rate with PI	4.2%	4.6%	5.1%

Note: MPCs refer to MPCs out of a $500 dollar stimulus check. "HANK: CS + OC" is our model in which we allow for skill heterogeneity and overconfidence, "Standard HANK" denotes a standard one-asset model, in which we abstract from heterogeneity in skills and overconfidence, "Standard HANK low wealth" is the same HANK model but with restricted liquidity to match the average MPC of "HANK: CS + OC". "... with PI" refers to the stationary equilibrium in the models with public insurance via low-income transfers.

We compare the stationary equilibrium effects of targeted transfers on the average MPC and HtM share in our model vs. in a standard rational one-asset HANK model. In addition, we also consider the rational HANK model in which we reduce the amount of wealth such that it produces the same average MPC in the absence of transfers as our model with skill and belief heterogeneity ("standard HANK, low wealth").[33]

In the two standard one-asset models without skill and belief heterogeneity, targeted transfers crowd-out self-insurance precautionary savings in the stationary equilibrium. Households correctly forecast their eligibility for transfer receipt during the next bad productivity shock, and those expecting to be eligible reduce their steady-state precautionary saving accordingly. This increases the average MPC from 0.164 to 0.215 (or from 0.032 to 0.045 in the model without the adjusted wealth level), and the HtM share from 21% to 24%. The crowding-out of private insurance can also be seen by the relatively strong increase in the equilibrium real rate from 4% to 5.1%. This higher real rate is necessary to make households hold the supplied liquidity in equilibrium. Thus, targeted transfers as social insurance produce an economy with less precautionary savings, a higher share of HtM households, and a higher average MPC.

In our one-asset model with skill and belief heterogeneity, these crowding-out effects are dampened because overconfident households underpredict their probability of reaching a low-productivity state in which they receive a transfer. The average MPC even slightly decreases from 0.164 to 0.158, while the share of HtM households remains practically unchanged (from 25.2% to

[33] As discussed in Section 4.2, calibrating the rational one-asset HANK model to generate the same average MPC as our model with heterogeneity in skills and beliefs, requires an average wealth to average annual income ratio of about 60% in the rational model, compared to 410% as in the data and in our baseline model with overconfidence.

26%). The increase in the real interest rate is also substantially weaker than in the model without skill and belief heterogeneity: the real rate in our model increases only from 4% to 4.2%.

5.2.2 Targeted transfers as business cycle instrument

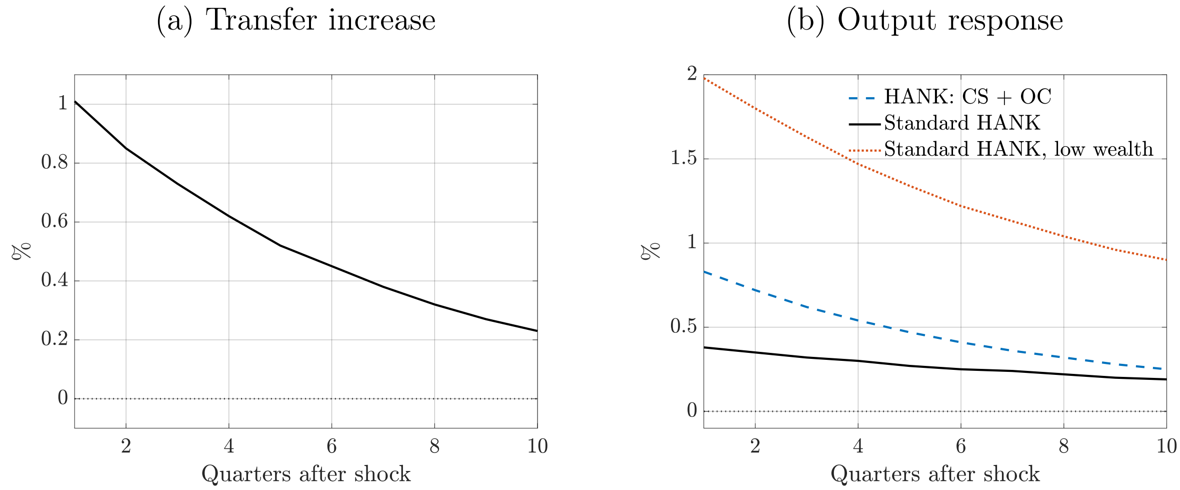
We now analyze the effectiveness of targeted transfers as a short-run stabilization policy. In particular, we model these transfers as direct cash transfers to the bottom 25% households in terms of their income. The aggregate expenditures for all transfers are 1 percent of steady state output on impact, with a persistence of 0.8. We consider the same three models as in the section before: our baseline HANK model with heterogeneity in cognitive skills and overconfidence, the standard rational HANK model without skill heterogeneity and in which all households hold rational beliefs about their idiosyncratic risk, and the the standard HANK model in which we restrict the amount of wealth in order to generate the same average MPC as our model with heterogeneity in skills and beliefs.

Panel (a) in Figure 4 shows the exogenous path of the transfers (relative to steady state output), and panel (b) shows the output deviations from steady state in percentage terms for the three different economies. Our model with heterogeneity in cognitive skills and overconfidence (blue-dashed line) predicts an output response that is about twice as large as in its rational counterpart (black-solid line). The key difference of these two models is that our model features a substantially higher average MPC and thus, leads to a higher consumption response to the transfers.

Compared to the standard HANK model with low wealth which has the same average MPC (orange-dotted line), however, our model predicts a substantially smaller output response. The reason for that is twofold. First, there is a lower correlation in our model between income and MPCs compared to the standard HANK with low wealth as in our model a key predictor of being HtM is whether a household is overconfident or not. As such, transfer recipients—those with lower incomes—have a smaller average MPC in our baseline model. Figure 5 illustrates that the income-HtM distribution of our model fits the data quite well and, importantly, much better than the standard HANK model with low wealth. It shows the share of HtM households for the four income quartiles in the data (black dots), our model (blue dots) and the standard HANK model with low wealth (orange dots). The standard HANK model substantially overestimates the share of HtM households in the lowest quartile, which are the households that receive the income transfers. Thus, the standard HANK model likely overpredicts the stimulative effects of these transfers. Our model with heterogeneity in cognitive skills and overconfidence, in contrast, almost perfectly matches the share of HtM households in the lowest quartile (and also the second quartile) and, thus, seems to be a good fit to analyze the effectiveness of targeted transfers.

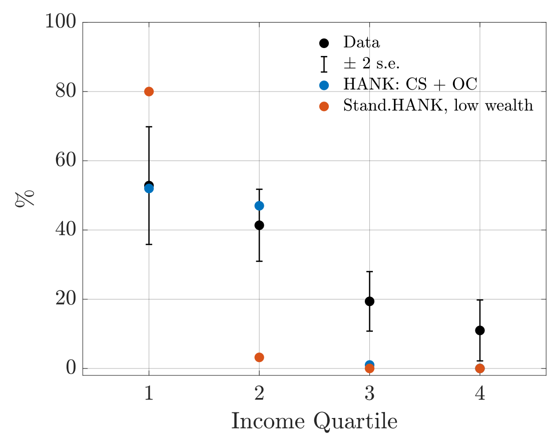
The second mechanism that tends to dampen the effects of targeted transfers in our model is that temporarily higher transfers relax households' precautionary-savings motive which however is muted in our model compared to rational models. As long as transfers are increased, households'

Figure 4: Targeted Transfer Shocks



Note: This figure shows the effects of a positive transfer shock ε_t^{TT} (left panel) on total output (right panel). Both are expressed in percentage deviations from steady state output. The blue-dashed line shows the response of our baseline model featuring permanent differences in skills and overconfidence. The black-solid line and the red-dotted line show the responses in the standard HANK model with rational expectations and without permanent skill differences for different wealth levels. The red-dotted line shows the case in which the wealth level in this economy is adjusted such that it features the same aggregate MPC as our model with skill and belief heterogeneity.

Figure 5: Distribution of HtM along the income distribution



Note: This figure shows the share of hand-to-mouth households along the income distribution in our baseline model with overconfidence and in the standard HANK model recalibrated to match the average MPC of our baseline model.

income risk is lower. Households therefore reduce their savings and consume more, leading to an increase in total output.[34] Yet, this second effect is dampened once we account for the high share of overconfident households in the data, because overconfident households underestimate their risk of reaching low income states which are the states in which higher transfers are paid out. Thus,

[34] See for example Bayer et al. (2023a) for an analysis of targeted transfers in a rational HANK model in which the relaxation of households' precautionary savings are an important driver of their high multipliers. In addition, Kekre (2023) analyzes temporary increases in unemployment benefits and Dengler and Gehrke (2022) analyzes increases in "short-time work" both of which can be seen as special versions of targeted transfers. Also in these cases, the relaxation of the precautionary savings motive significantly increases the effectiveness of these fiscal policies.

they *perceive* their precautionary savings motive to be less relaxed as rational households would do, such that they do not increase their spending by as much.[35] This second channel therefore offers a mechanism to the empirical findings in D'Acunto et al. (2023a) who show that cognitive constraints can limit the effectiveness of certain policies.

6 Conclusion

We analyze the implications of household heterogeneity in cognitive skills and perceptions thereof for households' financial situations and their savings behavior. Using U.S. micro level data, we find that lower-skilled households systematically over-estimate their skills and are persistently overly optimistic about their future financial situations. Additionally, they are substantially more likely to be hand-to-mouth.

Introducing permanent skill heterogeneity and overconfidence into a HANK model allows us to replicate these empirical patterns. What is more, our model can resolve heretofore seemingly intrinsic tensions in HANK models. Unlike other models, our one-asset HANK model can simultaneously match consensus estimates of both the average MPC and the average wealth level. Our two-asset HANK model matches the data with a lower, and perhaps more empirically realistic, liquidity premium than required in other models.

We show that accounting for the underlying reason why some households are systematically more likely to be at or close to the borrowing constraint matters greatly for fiscal policy—both positively and normatively. Normatively, we show that the optimal amount of government debt is substantially lower than in the standard model, because issuing more debt and thereby increasing the supply of the liquid asset is less effective in bringing households away from the borrowing constraint and in reducing wealth inequality than in standard one-asset HANK models. Positively, we find that targeted transfers to low-income households have weaker crowding-out effects of households' self insurance, as overconfident households do not privately insure themselves against these low-income states to begin with. Temporarily increasing income-targeted transfers to stimulate the economy is less effective in our model once we compare it to rational models with the same average MPC as the effects of the induced relaxation of households' precautionary-savings motive are weaker and the average MPC of transfer recipients is lower.

[35] The relaxation of the precautionary savings motive of households is also an important driver for the aggregate effects of these transfers in the standard HANK model. Yet, as the MPCs are so low in that model, it still has a significantly smaller effect on aggregate output.

References

Agarwal, S. and B. Mazumder (2013): "Cognitive abilities and household financial decision making," *American Economic Journal: Applied Economics*, 5, 193–207.

Aguiar, M. A., M. Bils, and C. Boar (2021): "Who are the Hand-to-Mouth?" .

Aiyagari, S. R. (1994): "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 109, 659–684.

Aiyagari, S. R. and E. R. McGrattan (1998): "The optimum quantity of debt," *Journal of Monetary Economics*, 42, 447–469.

Almgren, M., J.-E. Gallegos, J. Kramer, and R. Lima (2022): "Monetary policy and liquidity constraints: Evidence from the euro area," *American Economic Journal: Macroeconomics*, 14, 309–340.

Andreoni, J. and C. Sprenger (2012): "Estimating time preferences from convex budgets," *American Economic Review*, 102, 3333–3356.

Andreou, E., M. Demetriadou, and A. Tryphonides (2023): "The Cross Section of Household Preferences and the Marginal Propensity to Consume: Evidence from high-frequency data," .

Angeletos, G.-M. and Z. Huo (2021): "Myopia and anchoring," *American Economic Review*, 111, 1166–1200.

Auclert, A. (2019): "Monetary policy and the redistribution channel," *American Economic Review*, 109, 2333–67.

Auclert, A., B. Bardóczy, and M. Rognlie (2021): "MPCs, MPEs, and multipliers: A trilemma for New Keynesian models," *The Review of Economics and Statistics*, 1–41.

Auclert, A., M. Rognlie, and L. Straub (2020): "Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model." .

——— (2023): "The intertemporal keynesian cross," .

Balleer, A., G. Duernecker, S. Forstner, and J. Goensch (2022): "The Effects of Biased Labor Market Expectations on Consumption, Wealth Inequality, and Welfare," .

Banks, J. and Z. Oldfield (2007): "Understanding pensions: Cognitive function, numerical ability and retirement saving," *Fiscal studies*, 28, 143–170.

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro (1997): "Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study," *The Quarterly Journal of Economics*, 112, 537–579.

Bayer, C., B. Born, and R. Luetticke (2022): "The liquidity channel of fiscal policy," *Journal of Monetary Economics.*

——— (forthcoming): "Shocks, frictions, and inequality in US business cycles," *American Economic Review.*

Bayer, C., B. Born, R. Luetticke, and G. Müller (2023a): "The Coronavirus Stimulus

Package: How large is the transfer multiplier?" *The Economic Journal*.

Bayer, C., A. Kriwoluzky, G. Mueller, and F. Seyrich (2023b): "Das Sein Bestimmt das Bewußtsein - Fiscal attitudes according to HANK^2," .

Bayer, C., R. Lütticke, L. Pham-Dao, and V. Tjaden (2019): "Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk," *Econometrica*, 87, 255–290.

Benigno, P. and L. Paciello (2014): "Monetary policy, doubts and asset prices," *Journal of Monetary Economics*, 64, 85–98.

Bewley, T. (1986): "Stationary monetary equilibrium with a continuum of independently fluctuating consumers," *Contributions to mathematical economics in honor of Gérard Debreu*, 79.

Bhandari, A., D. Evans, M. Golosov, and T. J. Sargent (2021): "Inequality, Business Cycles, and Monetary-Fiscal Policy," *Econometrica*, 89, 2559–2599.

Bilbiie, F. O. (2021): "Monetary policy and heterogeneity: An analytical framework," .

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in macroeconomic expectations," *American Economic Review*, 110, 2748–82.

Broer, T., N.-J. Harbo Hansen, P. Krusell, and E. Öberg (2020): "The New Keynesian transmission mechanism: A heterogeneous-agent perspective," *The Review of Economic Studies*, 87, 77–101.

Caballero, R. J. and A. Simsek (2020): "Prudential monetary policy," .

Caplin, A., V. Gregory, E. Lee, S. Leth-Petersen, and J. Sæverud (2023): "Subjective Earnings Risk," .

Chapman, J., M. Dean, P. Ortoleva, E. Snowberg, and C. Camerer (2023): "Econographics," *Journal of Political Economy Microeconomics*, 1, 115–161.

Dávila, E. and A. Schaab (2022): "Welfare Assessments with Heterogeneous Individuals," .

——— (2023): "Optimal monetary policy with heterogeneous agents: A timeless ramsey approach," .

Dengler, T. and B. Gehrke (2022): "Short-time work and precautionary savings," .

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010): "Are risk aversion and impatience related to cognitive ability?" *American Economic Review*, 100, 1238–1260.

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber (2019): "Cognitive abilities and inflation expectations," in *AEA Papers and Proceedings*, vol. 109, 562–66.

——— (2023a): "Cognitive constraints and economic incentives," .

——— (2023b): "IQ, expectations, and choice," *Review of Economic Studies*.

Ehrlinger, J., K. Johnson, M. Banner, D. Dunning, and J. Kruger (2008): "Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent," *Organizational behavior and human decision processes*, 105, 98–121.

Erceg, C. J., D. W. Henderson, and A. T. Levin (2000): "Optimal monetary policy with staggered wage and price contracts," *Journal of monetary Economics*, 46, 281–313.

Fagereng, A., M. B. Holm, and G. J. Natvik (2021): "MPC heterogeneity and household balance sheets," *American Economic Journal: Macroeconomics*, 13, 1–54.

Farhi, E. and I. Werning (2019): "Monetary policy, bounded rationality, and incomplete markets," *American Economic Review*, 109, 3887–3928.

Gabaix, X. (2014): "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 129, 1661–1710.

——— (2020): "A behavioral New Keynesian model," *American Economic Review*, 110, 2271–2327.

Gillen, B., E. Snowberg, and L. Yariv (2019): "Experimenting with measurement error: Techniques with applications to the caltech cohort study," *Journal of Political Economy*, 127, 1826–1863.

Guerreiro, J. (2023): "Belief Disagreement and Business Cycles," .

Hagedorn, M., J. Luo, I. Manovskii, and K. Mitman (2019): "Forward guidance," *Journal of Monetary Economics*, 102, 1–23.

Hanushek, E. A. and L. Woessmann (2008): "The role of cognitive skills in economic development," *Journal of economic literature*, 46, 607–668.

Havranek, T. and A. Sokolova (2020): "Do consumers really follow a rule of thumb? Three thousand estimates from 144 studies say "probably not"," *Review of Economic Dynamics*, 35, 97–122.

Heidhues, P., B. Kőszegi, and P. Strack (2018): "Unrealistic expectations and misguided learning," *Econometrica*, 86, 1159–1214.

Holm, M. B., P. Paul, and A. Tischbirek (2021): "The transmission of monetary policy under the microscope," *Journal of Political Economy*, 129, 2861–2904.

Huffman, D., C. Raymond, and J. Shvets (2022): "Persistent overconfidence and biased memory: Evidence from managers," *American Economic Review*, 112, 3141–75.

Huggett, M. (1993): "The risk-free rate in heterogeneous-agent incomplete-insurance economies," *Journal of economic Dynamics and Control*, 17, 953–969.

Ilut, C. and R. Valchev (2023): "Economic agents as imperfect problem solvers," *The Quarterly Journal of Economics*, 138, 313–362.

Jappelli, T. and L. Pistaferri (2010): "The consumption response to income changes," *Annu. Rev. Econ.*, 2, 479–506.

——— (2014): "Fiscal policy and MPC heterogeneity," *American Economic Journal: Macroeconomics*, 6, 107–136.

Jensen, A. R. (1998): "The g factor and the design of education," *Intelligence, instruction, and assessment: Theory into practice*, 111–131.

Jordà, Ò., K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor (2019): "The rate of return on everything, 1870–2015," *The Quarterly Journal of Economics*, 134, 1225–1298.

Kaplan, G., B. Moll, and G. L. Violante (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

Kaplan, G. and G. L. Violante (2014): "A model of the consumption response to fiscal stimulus payments," *Econometrica*, 82, 1199–1239.

——— (2022): "The marginal propensity to consume in heterogeneous agent models," *Annual Review of Economics*, 14, 747–775.

Kekre, R. (2023): "Unemployment Insurance in Macroeconomic Stabilization," *Review of Economic Studies*.

Kekre, R. and M. Lenel (2022): "Monetary policy, redistribution, and risk premia," *Econometrica*, 90, 2249–2282.

Krueger, D., K. Mitman, and F. Perri (2016): "Macroeconomics and household heterogeneity," in *Handbook of Macroeconomics*, Elsevier, vol. 2, 843–921.

Laibson, D., P. Maxted, and B. Moll (2021): "Present bias amplifies the household balance-sheet channels of macroeconomic policy," .

Lian, C. (forthcoming): "Mistakes in future consumption, high MPCs now," *American Economic Review: Insights*.

Luetticke, R. (2021): "Transmission of monetary policy with heterogeneity in household portfolios," *American Economic Journal: Macroeconomics*, 13, 1–25.

Lusardi, A. and O. S. Mitchell (2014): "The economic importance of financial literacy: Theory and evidence," *American Economic Journal: Journal of Economic Literature*, 52, 5–44.

MacLeod, C. M. (1991): "Half a century of research on the Stroop effect: an integrative review." *Psychological bulletin*, 109, 163.

McArdle, J. J., G. G. Fisher, and K. M. Kadlec (2007): "Latent variable analyses of age trends of cognition in the Health and Retirement Study, 1992-2004." *Psychology and aging*, 22, 525.

McClung, N. and T. Nighswander (2021): "Belief Heterogeneity and the Income-Wealth Relationship," .

McKay, A., E. Nakamura, and J. Steinsson (2016): "The power of forward guidance revisited," *The American Economic Review*, 106, 3133–3158.

McKay, A. and C. Wolf (2023): "Optimal policy rules in hank," .

Miyake, A. and N. P. Friedman (2012): "The nature and organization of individual differences in executive functions: Four general conclusions," *Current directions in psychological science*, 21, 8–14.

Moschini, E., G. Raveendranathan, and M. Xu (2023): "Over-optimism About Graduation and College Financial Aid," .

Mueller, A. I., J. Spinnewijn, and G. Topa (2021): "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias," *American Economic Review*, 111, 324–363.

Patterson, C. (2023): "The matching multiplier and the amplification of recessions," *American*

Economic Rewview.

Pfäuti, O. and F. Seyrich (2022): "A behavioral heterogeneous agent new keynesian model," .

Rouwenhorst, K. G. (1995): "10 Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of business cycle research*, Princeton University Press, 294–330.

Rozsypal, F. and K. Schlafmann (2023): "Overpersistence bias in individual income expectations and its aggregate implications," *American Economic Journal: Macroeconomics.*

Seidl, H. and F. Seyrich (2021): "Unconventional fiscal policy in HANK," *DIW Berlin Discussion Paper.*

Sergeyev, D., C. Lian, and Y. Gorodnichenko (2023): "The Economics of Financial Stress," .

Smirnov, D. (2022): "Optimal monetary policy in hank," *Working paper.*

Solon, G., S. J. Haider, and J. M. Wooldridge (2015): "What are we weighting for?" *Journal of Human resources*, 50, 301–316.

Souleles, N. S. (2004): "Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the Michigan consumer sentiment surveys," *Journal of Money, Credit and Banking*, 39–72.

Stango, V. and J. Zinman (2022): "Behavioral biases are temporally stable," .

——— (2023): "We Are All Behavioural, More, or Less: A Taxonomy of Consumer Decision-Making," *The Review of Economic Studies*, 90, 1470–1498.

Wang, T. (2023): "Perceived versus calibrated income risks in heterogeneous-agent consumption models," .

Werning, I. (2015): "Incomplete markets and aggregate demand," .

Wolf, C. (2021): "Interest Rate Cuts vs. Stimulus Payments: An Equivalence Result," .

Woodford, M. (1990): "Public debt as private liquidity," *The American Economic Review*, 80, 382–388.

——— (2013): "Macroeconomic analysis without the rational expectations hypothesis," *Annu. Rev. Econ.*, 5, 303–346.

——— (2019): "Monetary policy analysis when planning horizons are finite," *NBER Macroeconomics Annual*, 33, 1–50.

Yang, Y. (2022): "Redistributive inflation and optimal monetary policy," *Working paper.*

Zimmermann, F. (2020): "The dynamics of motivated beliefs," *American Economic Review*, 110, 337–363.

A Additional Tables and Results

Table A1: Regression results: income forecasts and financial condition forecasts

	Nom. income, unw. (1)	Nom. income, w. (2)	Real income, unw. (3)	Real income, w. (4)
Coefficient	0.0049	0.0049	0.0057	0.0055
s.e.	(0.0001)	(0.0002)	(0.0001)	(0.0002)
N	10,745	10,745	10,743	10,743

Notes: This table shows the regression results when regressing our measure of optimists according to expected future financial conditions on different measures of income forecasts. The regression includes a constant. Columns (1) and (3) show the results for the unweighted sample, whereas columns (2) and (4) show the results when using sampling weights. A point estimate of 0.005 says our optimism measure using subjective financial conditions is 0.5pp higher for an increase of 1pp in the forecasted likelihood of a (nominal or real) income increase.

Table A2: Household financial condition forecasts and forecast errors tilt optimistic

Panel A. All forecasts, unweighted	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.09	0.13	0.04	0.27
Same	0.06	0.44	0.10	0.61
Worse	0.01	0.05	0.07	0.12
Total	0.16	0.63	0.21	1
Panel B. July 2009 & 2010, unweighted	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.06	0.16	0.05	0.28
Same	0.05	0.40	0.15	0.60
Worse	0.01	0.05	0.07	0.12
Total	0.12	0.61	0.27	1
Panel C. July 2009 & 2010, weighted	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.07	0.18	0.05	0.30
Same	0.04	0.38	0.14	0.56
Worse	0.01	0.07	0.06	0.14
Total	0.12	0.63	0.25	1
Panel D. January 2015 & 2016, unweighted	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.10	0.14	0.04	0.28
Same	0.06	0.47	0.08	0.61
Worse	0.01	0.05	0.06	0.12
Total	0.17	0.66	0.18	1
Panel E. January 2015 & 2016, weighted	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.11	0.13	0.03	0.27
Same	0.05	0.50	0.08	0.63
Worse	0.01	0.04	0.05	0.10
Total	0.17	0.67	0.16	1

Note: Cells report sample proportions. Forecasts: "Now looking ahead - do you think that a year from now you will be better off financially, or worse off, or just about the same as now?" Response options: Will be better off/About the same/Will be worse off. Realizations: "We are interested in how people are getting along financially these days. Would you say that you are better off or worse off financially than you were a year ago?" Response options: Better off/About the same/Worse off. Weighted estimates use sampling probabilities from the realization survey(s), which are correlated 0.90 and 0.93 with the weight from the paired forecast survey. Sample size is 17,266 in Panel A, 1,679 in Panels B and C, and 1,882 in Panels D and E.

Table A3: Household financial condition forecast errors are persistent

	Forecast error this survey			
FCE previous survey	Optimist	Realist	Pessimist	Total
Optimist	0.08	0.10	0.00	0.18
Realist	0.07	0.65	0.03	0.75
Pessimist	0.01	0.04	0.02	0.06
Total	0.16	0.79	0.05	1

Note: Sample is 6,590 forecast error pairs from 2,964 panelists. Sample is smaller here than in Appendix Table A2 because here we require >=2 forecast-realization pairs per panelist and only include realizations of "about the same", to allow for capturing forecast errors in either direction.

Table A4: Household financial condition forecast learning?

Panel A. First forecast - realization pair	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.09	0.16	0.06	0.31
Same	0.05	0.40	0.12	0.58
Worse	0.01	0.05	0.06	0.12
Total	0.15	0.61	0.23	1
Panel B. Last forecast - realization pair	Realization this year			
Forecast last year	Better	Same	Worse	Total
Better	0.10	0.13	0.04	0.27
Same	0.06	0.46	0.09	0.61
Worse	0.01	0.05	0.06	0.12
Total	0.17	0.64	0.19	1

Note: Sample only considers forecast - realization pairs with multiple pairs, resulting in 2,964 panelists.

Table A5: Pairwise correlations between persistent overconfidence about cognitive skills and persistent optimistic forecast errors

	1 = oc both rounds		oc percentile rank		Mean(row var)	
	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted
	(1)	(2)	(3)	(4)	(5)	(6)
1=(Consec. opt. FEs)	0.163	0.082	0.107	0.029	0.147	0.124
s.e.	0.095	0.123	0.056	0.066		
N	409	409	409	409		
1=(Prop. opt. FEs $\geq$ 0.5)	0.219	0.197	0.155	0.149	0.130	0.140
s.e.	0.097	0.145	0.055	0.091		
N	409	409	409	409		
Prop. opt. FEs	0.096	0.108	0.134	0.159	0.174	0.181
s.e.	0.056	0.084	0.057	0.073		
N	409	409	409	409		

Note: Overconfidence re: relative performance in a cognitive skills test (see Section 2 for details). Forecast errors re: household financial condition (see Appendix Table A2 for details). Weighted estimates use the mean of each panelist's: (sample probably weight from the last SZ module, mean sampling weight across the survey(s) with the realization component of the forecast error(s) used here). In Columns (3) and (4), we use Obviously Related Instrumental Variables to account for measurement error by having the two measurements of o/c rank (taken in 2014 and 2017) instrument for each other (Gillen et al. (2019), Stango and Zinman (2023)). We do not take the same approach to the overconfidence indicator in Columns (1) and (2), because measurement error-IV does not work well on misclassification error. Fully non-IV correlations estimated using tetrachoric or Pearson.

Table A6: Weighted pairwise correlations between cognitive skills and income

	Cognitive skills undimensional summstats		
	1st cf score (1)	1st principal component (2)	rowvar mean
Household income	0.522	0.562	56351
s.e. or s.d.	0.053	0.055	43618
N	766	662	1580
ln(Household income)	0.555	0.568	10.548
s.e. or s.d.	0.062	0.068	1.502
N	803	697	1637
Household income category	0.526	0.540	10.839
s.e. or s.d.	0.044	0.051	4.317
N	842	732	1686

Note: Weighted ORIV correlations. Column variables are cross-sectional percentiles. Level income drops obs. in top 5 percentiles. Categorical income based on ALP's standard income elicitation, in which respondents choose among 17 bins.

Table A7: Persistent overconfidence: Population share, and correlations with cognitive skills

	1 = oc both rounds		oc percentile rank	
	Unweighted	Weighted	Unweighted	Weighted
	(1)	(2)	(3)	(4)
Population share	0.340	0.377		
s.e.	0.017	0.035		
N	817	817		
Cognitive skill measures				
Summary: 1st principal component	-0.546	-0.542	-0.818	-0.830
s.e.	0.030	0.045	0.032	0.049
N	733	733	733	733
Component: Fluid intelligence	-0.718	-0.734	-1.049	-1.065
s.e.	0.026	0.047	0.026	0.055
N	817	817	817	817
Component: Numeracy	-0.362	-0.453	-0.573	-0.656
s.e.	0.040	0.068	0.046	0.077
N	798	798	798	798
Component: Financial literacy	-0.321	-0.242	-0.467	-0.362
s.e.	0.038	0.087	0.041	0.087
N	813	813	813	813
Component: Executive function	-0.316	-0.407	-0.444	-0.600
s.e.	0.045	0.072	0.052	0.090
N	749	749	749	749

Note: Overconfidence re: relative performance in a cognitive skills test (see Section 2 for details). All cognitive skills measures are percentile ranks. Cognitive skills summary measure is the first common factor of each of the component measures shown in the table (see Stango and Zinman (2023) for details on component measures). Weighted estimates use the sampling probability for the last SZ module. All cognitive skills measures, and overconfidence percentile rank, use Obviously Related Instrumental Variables to account for measurement error by having the two rank measures (taken in 2014 and 2017) instrument for each other (Gillen et al. (2019), Stango and Zinman (2023)). We do not take the same approach to the overconfidence indicator in Columns (1) and (2), because measurement error-IV does not work well on misclassification error.

Table A8: Pairwise correlations between persistent optimistic forecast errors and patience and risk aversion

	Patience		Risk aversion			
	Unw.	Weighted	Unw.	Weighted	Unw.	Weighted
	(1)	(2)	(3)	(4)	(5)	(6)
1=(Prop. opt. FEs $\geq$ 0.5)	-0.09	-0.10	-0.10	-0.20	-0.06	-0.12
s.e.	(0.079)	(0.143)	(0.063)	(0.096)	(0.067)	(0.098)
N	340	340	352	352	353	353
Prop. opt. FEs	-0.11	-0.12	-0.12	-0.25	-0.06	-0.11
s.e.	(0.081)	(0.160)	(0.063)	(0.090)	(0.067)	(0.104)
N	340	340	352	352	353	353

Notes: Weighted estimates use sampling probability from the last SZ module. Forecast errors re: household financial condition and are defined at the household level over multiple observations (see Appendix Table A2 for details). Patience is the average savings rate across 24 convex time budget choices (Andreoni and Sprenger, 2012). Risk aversion in Columns (3) and (4) is based on the Dohmen et al. (2010) financial risk-taking scale, and in Columns (5) and (6) on the Barsky et al. (1997) lifetime income gamble elicitation. We use Obviously Related Instrumental Variables to account for measurement error in the column variables by using the two measures of each (taken in 2014 and 2017) to instrument for each other (Gillen et al., 2019; Stango and Zinman, 2023).

Table A9: Pairwise correlations between persistent HtM measures and patience and risk aversion

	Patience		Risk aversion			
	Unw.	Wtd.	Unw.	Wtd.	Unw.	Wtd.
	(1)	(2)	(3)	(4)	(5)	(6)
1=(Severe financial distress)	-0.014	-0.081	0.107	0.029	0.036	0.077
s.e.	(0.057)	(0.143)	(0.042)	(0.091)	(0.049)	(0.123)
N	780	780	818	818	832	832
1=(Low net worth)	-0.025	-0.073	0.057	0.080	0.136	0.032
s.e.	(0.058)	(0.098)	(0.042)	(0.074)	(0.050)	(0.090)
N	734	734	765	765	778	778
1=(paycheck-to-paycheck c. 2012)	0.062	0.377	0.010	0.069	0.048	-0.157
s.e.	(0.100)	(0.167)	(0.073)	(0.164)	(0.088)	(0.311)
N	233	233	256	256	260	260
paycheck-to-paycheck, COVID era	-0.126	-0.014	0.084	0.051	0.130	0.007
s.e.	(0.073)	(0.120)	(0.051)	(0.075)	(0.057)	(0.098)
N	493	493	516	516	519	519
1=(Lacks prec. saving in 2012 & 2018)	-0.218	-0.186	0.114	0.051	0.068	-0.078
s.e.	(0.083)	(0.127)	(0.070)	(0.114)	(0.077)	(0.140)
N	254	254	264	264	269	269
Difficult covering $2k$ emerg. expenses	-0.154	-0.039	0.136	0.146	0.108	0.133
s.e.	(0.065)	(0.117)	(0.051)	(0.078)	(0.058)	(0.108)
N	462	462	487	487	491	491

Note: Weighted estimates use the sampling probability for the last SZ module. Patience is the average savings rate across 24 convex time budget choices (Andreoni and Sprenger, 2012). Risk aversion in Columns 3 and 4 is based on the Dohmen et al. (2010) financial risk-taking scale, and in Columns 5 and 6 on the Barsky et al. (1997) lifetime income gamble elicitation. We use Obviously Related Instrumental Variables to account for measurement error in the column variables, and in overconfidence percentile rank, by using the two measures of each (taken in 2014 and 2017) to instrument for each other (Gillen et al., 2019; Stango and Zinman, 2023). Severe financial distress=1 if panelist reported any of four events (forced move, late payments, hunger, foregone medical care) in past 12 months in both SZ modules. Low wealth=1 if net worth<1/2 total monthly household income in both SZ modules. Paycheck-to-paycheck c. 2012 survey =1 if panelist strongly agrees with: 'I live from paycheck to paycheck'. Paycheck-to-paycheck for COVID era is the proportion of up to 9 surveys, from May 2020-July 2022, where panelist responds 'Very difficult' or 'Somewhat difficult' to 'In the past month, how difficult has it been for you to cover your expenses and pay all your bills?' OR if on the followup q. 'Suppose now you have an emergency expense that costs $400. Based on your current financial situation, how would you pay this expense?' they report one or more expensive options: credit card revolving, small-dollar credit, wouldn't be able to pay for it. Lacks precautionary savings=1 if panelist does not have emerg/rainy day funds set side to cover 3-months' expenses. Difficulty covering expense is the proportion of 3 surveys from 2011, 2012, and 2018 where panelist does not express the highest confidence or certainty that they could cover an unexpected $2,000 need arising in the next month. Paycheck-to-paycheck c. 2012 survey =1 if panelist strongly agrees with: 'I live from paycheck to paycheck'. Paycheck-to-paycheck for COVID era is the proportion of up to 9 surveys, from May 2020-July 2022, where panelist responds 'Very difficult' or 'Somewhat difficult' to 'In the past month, how difficult has it been for you to cover your expenses and pay all your bills?' OR if on the followup q. 'Suppose now you have an emergency expense that costs $400. Based on your current financial situation, how would you pay this expense?' they report one or more expensive options: credit card revolving, small-dollar credit, wouldn't be able to pay for it. Lacks precautionary savings=1 if panelist does not have emerg/rainy day funds set side to cover 3-months' expenses. Difficulty covering expense is the proportion of 3 surveys from 2011, 2012, and 2018 where panelist does not express the highest confidence or certainty that they could cover an unexpected $2,000 need arising in the next month.

Table A10: Correlations between overconfidence and patience and risk aversion

	Patience		Risk Aversion			
	Unwtd.	Weighted	Unwtd.	Weighted	Unwtd.	Weighted
	(1)	(2)	(3)	(4)	(5)	(6)
1=Oc both rounds	0.04	-0.01	-0.08	-0.20	0.16	0.24
s.e.	(0.056)	(0.141)	(0.042)	(0.074)	(0.050)	(0.120)
N	758	758	813	813	807	807
Oc percentile rank	0.00	-0.01	-0.15	-0.32	0.24	0.31
s.e.	(0.066)	(0.118)	(0.049)	(0.079)	(0.056)	(0.116)
N	758	758	813	813	807	807

Notes: Weighted estimates use sampling probability from the last SZ module. Discrete measure of overconfidence defined as exhibiting above-median confidence in relative performance on a fluid intelligence test in both 2014 and 2017, see Section 2 for details. Patience is the average savings rate across 24 convex time buget choices (Andreoni and Sprenger, 2012). Risk aversion in Columns (3) and (4) is based on the Dohmen et al. (2010) financial risk-taking scale, and in Columns (5) and (6) on the Barsky et al. (1997) lifetime income gamble elicitation. We use Obviously Related Instrumental Variables to account for measurement error in the column variables, and in overconfidence percentile rank, by using the two measures of each (taken in 2014 and 2017) to instrument for each other (Gillen et al., 2019; Stango and Zinman, 2023).