Narrow and Short Beliefs in Macroeconomics with Heterogeneous Agents

Sebastián Guarda*

November 12, 2025

[Updated often, click here for the latest version.]

Abstract

This paper studies economies with rich heterogeneity and aggregate uncertainty, where agents observe only a subset of equilibrium objects and rely on finite lags of observables to form forecasts. These belief frictions give rise to a novel equilibrium notion in which the agents' state space is always finite-dimensional, as opposed to full information and rational expectations (FIRE) where individuals may need to keep track of unwieldy distributions. In addition to allowing for a global solution with standard recursive methods, I show how these belief frictions can be disciplined with micro data on expectations by analyzing whether forecast errors are systematically predictable. As an application, I study a neoclassical model with heterogeneous income and wealth, where agents cannot differentiate between the idiosyncratic and aggregate components of their income. Relative to FIRE, aggregate consumption underreacts and investment overreacts persistently to productivity shocks. These dynamics help reconcile the model with the observed business-cycle moments of consumption and investment.

I am most grateful to Gianluca Violante, Pietro Ortoleva, Karthik Sastry, Nobuhiro Kiyotaki, Ayşegül Sahin, John Grigsby, Aleksei Oskolkov, Richard Rogerson, Jonathan Payne, Moritz Lenel, Michal Kolesár, and many participants of the Princeton Macroeconomics Workshop for helpful comments and suggestions.

*Princeton University, sguarda@princeton.edu

1 Introduction

How much information do people use when forming their beliefs about the economy? We must take a stance on this question when specifying models of the economy with forward-looking behavior, and we commonly assume that economic agents have all of the information available and they use it correctly. This is formalized in the full information and rational expectations (FIRE) equilibrium notion, which has become the canonical solution concept in macroeconomics due to many desirable properties. However, these informational requirements become particularly salient in the case of heterogeneous-agent models, where the relevant information can include the entire distribution of wealth and income at every point in time, a very "wide" set of information. This width seems at odds with recent empirical evidence on expectations, which also points to people overweighting the most recent realizations of the data when producing forecasts: a neglect of "long" sets of relevant histories. Moreover, it poses the computational challenge of solving a Bellman equation with an infinite-dimensional state: the distribution. Can alternative theories of beliefs address the empirical evidence on expectations, allow for feasible computation, and preserve desirable characteristics of the rational expectations approach?

In this paper, I study one such theory by considering two deviations from FIRE in general equilibrium models with heterogeneous agents and aggregate uncertainty. The deviation from FI is *narrow beliefs:* agents do not observe all of the possibly relevant variables, but a finite subset of them, as in theories of incomplete information such as Lucas (1972). The deviation from RE is *short beliefs:* when agents produce forecasts, they can only condition them on a finite number of lags of the variables that they have observed, as in Fuster et al. (2012). For example, agents might not observe the entire distribution of wealth, but they might observe the prices with which they interact, and use a predictive model with a small number of their lags to forecast future prices.

Subject to these constraints, agents forecast their variables of interest as best as they can, to aid their decision-making. This is formalized by an equilibrium condition that ensures consistency between their perceived law of motion (a predictive model) and the equilibrium dynamics, when conditioning on the variables and lags allowed by their narrow and short beliefs. In the absence of further parametric assumptions on the perceived law of motion, the resulting novel equilibrium concept is a Nonparametric Restricted Perceptions Equilibrium (NRPE), which generalizes the parametric RPE first studied by Marcet and Sargent (1989).

There are two important implications of studying an NRPE in our models of interest.

First, with data on expectations, we can discipline how narrow and how short beliefs are. In particular, if forecast errors over variables of interest are systematically correlated with an observable, that variable should not be in the agent's information set—otherwise the agent could improve their forecasting. Similarly, if forecast errors are systematically correlated with a given lag length, we conclude that the agent's predictive model is restricted to use fewer lags. This gives empirical discipline to how much bounded rationality we should impose on the agents in our models, the most important difference between my approach and that of Krusell and Smith (1998), henceforth KS.

The second implication is that the agents' state space is reduced to a finite dimension, so we can solve their Bellman equation and the global solution of the model with standard recursive methods. This is the goal of Moll (2025), who advocates for alternatives to rational expectations. Crucially, this payoff only comes when adopting both narrow and short beliefs: having either in isolation would generically still result in an infinite-dimensional state. I provide a computational algorithm to obtain the nonparametric perceived law of motion, which allows us to get a global solution for the NRPE in heterogeneous-agent models with aggregate shocks.

I begin the exposition in Section 2 by going over simple examples where agents must form expectations over random variables with some exogenous distribution. This is meant to illustrate the effect of narrow and short beliefs on expectations and forecast errors as transparently as possible, while leaving the study of beliefs over endogenous equilibrium variables to the next section. Nevertheless, the intuition will still be useful to understand one of the core mechanisms of Section 4. Lastly, I explain the general concept underlying FIRE's restriction on forecast errors, how we can use it to discipline the amount of bounded rationality for agents in our models, and why we need to introduce the nonparametric (N) element into the RPE framework in order to achieve this discipline.

Section 3 then outlines the general framework for this paper. I start by defining a general environment that encompasses a large class of models with heterogeneous agents and aggregate uncertainty. I revisit the problem with FIRE in this environment: the distribution of idiosyncratic states is an infinite-dimensional variable in the agent's Bellman equation, which prevents the use of standard recursive methods. After discussing the desirable properties of FIRE, I introduce the NRPE with narrow and short beliefs and show how many of these properties are preserved: beliefs are pinned down by the environment granting robustness to the Lucas critique; agents are not systematically fooled in equilibrium, given their constraints; the researcher has no degrees of freedom—as long as the frictions are disciplined with data on expectations; and computational tractability.

In addition, I discuss some secondary benefits of the nonparametric aspect of the NRPE. Lastly, I provide a robust computational algorithm to solve this class of models.

As an application, in Section 4 I study the presence of narrow beliefs in a general equilibrium neoclassical model with heterogeneous households and aggregate productivity shocks. In this model, household income is composed of an exogenous idiosyncratic component (efficiency units of labor) and an aggregate component endogenously determined in equilibrium (the wage). I first show that, even if households do not observe aggregate productivity or any moment of the distribution, as long as they observe each of the components and the rental rate of capital, then the economy is practically identical to that of KS. The reason is that prices are informationally equivalent to capital and productivity, the KS states. Given that KS is a very good approximation to FIRE, this is also the case of this benchmark economy by extension.

Then, I take as given empirical evidence by Adams and Rojas (2024) that households have narrow beliefs in income components: when they experience income changes, they can't differentiate the role of each component. I show that in a standard calibration, this leads to a higher persistence of the aggregate component than the idiosyncratic component. Thus when households forecast their income, they do so with a weighted average of these persistences. The implication is that when an aggregate productivity shock hits, consumption underreacts and investment overreacts, which increases the stock of capital and thus the agent's labor income in a persistent fashion, amplifying the effect of the original shock. I find that this mechanism can fit the joint dynamics of aggregate consumption and investment in the data at a business cycle frequency. In particular, FIRE features excessively volatile consumption relative to investment, and excessive correlation between the two, while narrow beliefs match these moments well.

Overall, the aim of this paper is to leverage the predictability of forecast errors by the kinds of agents that populate our heterogeneous-agent models (e.g. households). This would not only allow us to solve these models, but also to study the micro and macro implications of a boundedly-rational theory of expectations that aligns more closely to the data.

Related Literature Empirical evidence for narrow beliefs comes from households who use local signals to form beliefs about aggregate outcomes, despite the wide availability of aggregate information that renders their local signals useless in comparison. Kuchler and Zafar (2019) show that local house prices influence household beliefs over national prices, and households who personally experience unemployment become more pessimistic about nationwide unemployment. D'Acunto et al. (2021) show that personal grocery prices significantly affect aggregate inflation expectations. These empirical findings are suggestive of incomplete information in the style of the seminal contribution of Lucas (1972), which is what I denote as narrow beliefs.

Short beliefs overweight recent realizations of the data, for which we can find empirical evidence in different forms. A broad literature has studied the particular case of experience effects, where there is explicit dependence on what people have lived through. The seminal papers of Malmendier and Nagel (2011, 2016) introduced the concept while showing how the experience of the Great Depression led to pessimism and decreased risk-taking, and how inflation experiences shape inflation expectations. The general concept is also present in over-extrapolation: Greenwood and Shleifer (2014) find that investor expectations of stock returns are positively correlated with recent returns, as well as the current stock price levels. A simple way to formalize short beliefs was studied by Fuster et al. (2012) where agents predict dividend growth with fewer lags than those relevant under the data generating process; in this paper I follow their general approach.

Using FIRE equilibria as a solution concept in heterogeneous-agent models pose two problems: (1) the amount of information that agents use seems implausible, and (2) the computation is technically challenging. The latter was prominently discussed by KS, who proposed an approximation algorithm to address it: replace the law of motion of the distribution with a perceived law of motion over a finite number of moments of the distribution. It turns out that in their model, this approach is remarkably successful—even when using just the first moment—in that it approximates the FIRE equilibrium extremely well. This success can be appreciated by the number of authors who use it as-is or its variations, such as in the study of workers with search frictions by Krusell et al. (2010), price-setting by Midrigan (2011), firms with heterogeneous productivity by Khan and Thomas (2013), and housing prices by Kaplan et al. (2020).

Even though KS do not explicitly advocate for their algorithm as solving the first issue of counterfactual information, a widely held view is that there is a bounded rationality interpretation to the equilibrium that they compute: even if it's not plausible that agents know and forecast the distribution of wealth, it is plausible that they know and forecast its first moment.[1] I argue that the problem with this interpretation is that at the time of writing, there has been no attempt to empirically discipline the extent of this kind of bounded rationality. In fact, under the KS approach a higher R^2 in the perceived law of motion is always interpreted positively, which is always achieved by including more predictors of the variable of interest. The tradeoff is purely computational: the inclusion

[1] Krusell and Smith (2006) discuss this interpretation, see also Reiter (2010) and Moll (2025).

of an additional state variable is costly, so it might be left out if the R^2 is high enough.[2]

In this sense, KS is part of the literature that attempts to find an approximate solution to the FIRE equilibrium. One strand of this literature computes the deterministic steady state, and uses perturbation methods around it, such as Reiter (2010), Boppart et al. (2018), and Auclert et al. (2021). A more recent strand goes back to the KS goal of finding the full global solution but armed with neural networks, see Azinovic et al. (2022), Fernández-Villaverde et al. (2023), and Gu et al. (2024). I attempt a fundamentally different exercise: to discipline the amount of bounded rationality by solving for the equilibrium that is consistent with observed expectations. This equilibrium may be quantitatively similar to FIRE in some cases where the belief frictions have no economic bite (see Section 4.2) but it also may be quite different (see Section 4.3).

Discerning between the kind (and magnitude) of information frictions that can be supported empirically has been the focus of a large literature that analyzes data on expectations. Mankiw et al. (2003) document disagreement in inflation expectations that is suggestive of sticky-information as in Mankiw and Reis (2002). More recently, the predictability of forecast errors has been analyzed by Coibion and Gorodnichenko (2015), Bordalo et al. (2020) and many others, and usually interpreted as over- or under-reaction to news.[3] Instead, this paper is closer in spirit to Molavi et al. (2024), who study agents whose forecasts are generated with a factor model that has less factors than those of the data generating process (also a factor model). They match the forecast error predictability observed in the data to that generated by their theory, and show how it helps to explain uncovered interest parity puzzles. In doing so, expectations data disciplines the extent of bounded rationality, conceptualized as the number of factors that agents actually use.

The specific equilibrium notion in this paper is an RPE, the first example of which can be found in Marcet and Sargent (1989) who studied it from a recursive least-squares learning perspective.[4] Indeed we can think of RPE generally as limit points of economies populated by agents who are estimating econometric models with their observables, as their number of observations goes to infinity. In the language of KS—who actually compute an RPE which approximates FIRE—the (restricted) perceived law of motion of the agents is consistent with the equilibrium law of motion. Work on this literature includes Hommes and Sorger (1998) and Hajdini (2023) who considered the special case

[2]See the JEDC special issue Den Haan et al. (2010) for an extensive discussion of accuracy.

[3]A useful survey and further analysis on the dynamic properties of forecast errors is provided by Angeletos et al. (2021).

[4]Interestingly, they called it a rational expectations equilibrium, despite pointing out that forecast errors are systematically related to lags of variables that agents observe.

of a consistent expectations equilibrium,[5] as well as Branch and Evans (2010) who called it an RPE. Moll (2025) points out its potential in reducing the dimensionality of the agent's Bellman equation in heterogeneous-agent models, and Baley and Turen (2023) use it for this purpose. Relative to this literature, this paper's contribution is twofold: first is to show how the implications on forecast errors can be helpful to pin down the specification of the perceived law of motion. Second, to the best of my knowledge, I provide the first nonparametric generalization of this equilibrium notion. This allows me to contribute a methodology that can discipline the researcher's choice of the agent's state space.

More generally, this paper fits into the literature that considers a broad range of deviations from FIRE in general equilibrium. One of the most prominent approaches is rational inattention as pioneered by Sims (2003), reviewed by Maćkowiak et al. (2023), and studied in heterogeneous-agent models with aggregate shocks in Broer et al. (2021). Inattention can be a source of narrow beliefs: if households observe unemployment or inflation but don't pay them any attention, their beliefs will be the same as if they did not observe them in the first place. So through the lens of narrow beliefs, incomplete information and total inattention are isomorphic.[6] In turn, short beliefs usually imply over-extrapolation as in Fuster et al. (2012), sharing similarities with the diagnostic expectations framework of Bordalo et al. (2022) or memory constraints as in Nagel and Xu (2022). Additionally, the framework of Molavi (2024) provides an alternative way to think about both narrow and short beliefs in conjunction, for representative agent models.

Lastly, the application under study is narrow beliefs by households who cannot differentiate between the idiosyncratic and aggregate components of their income. This particular information structure has been studied by Pischke (1995) and Ludvigson and Michaelides (2001) in endowment economies with constant interest rates, finding that narrow beliefs help in jointly explaining the excess smoothness and excess sensitivity of aggregate consumption. More recently Adams and Rojas (2024) show empirical evidence in favor of narrow beliefs, and study its implications in small open economies. This paper contributes a general equilibrium analysis of the implications of narrow beliefs in a production economy, which provides an important amplification channel.

Layout The rest of the paper proceeds as follows. Section 2 introduces narrow and short beliefs with simple examples, and then conceptualizes their relation with the

[5]This equilibrium prioritizes the intertemporal misspecification, but the general concept is the same.

[6]This interpretation of total inattention is closer to the Gabaix (2014, 2019) formulation of sparse inattention.

predictability of forecast errors. Section 3 outlines a general framework for a class of heterogeneous-agent models, in order to introduce the main equilibrium definition: the NRPE with narrow and short beliefs. Section 4 goes over an application of narrow beliefs over income components, in a neoclassical model with uninsurable idiosyncratic income risk. Section 5 concludes.

2 Simple Examples and Implications

This section goes over simple examples of narrow and short beliefs, in order to introduce the concepts and fix ideas. In these examples, random variables have a known and exogenous distribution in order to distill the intuition transparently before pursuing the objective of this paper: to apply these frictions to endogenous variables that don't follow simple processes in equilibrium. In the final subsection, I go over the general principle that underlies the results found in the examples.

2.1 Narrow Beliefs

Consider an agent who is interested in predicting a variable of interest y, which is determined by two jointly normal random variables x and z:

$$y = x + z + v,$$

where v is inherently unobservable white noise. x and z are jointly normally-distributed, with variances σ_x^2 and σ_z^2 respectively, and covariance σ_{xz}^2. Full information in this case would mean that the agent can observe both x and z before making a prediction of y. In this case, the forecast errors of an agent with rational expectations will be given by v, and they will be unpredictable.

Narrow beliefs are a simple form of incomplete information: the agent may only observe one of these variables before forecasting. For example, if an agent with rational expectations only observes x before having to predict y, then their forecast is given by:

$$\hat{y} := \mathbb{E}\left[y \mid x\right] = x + \mathbb{E}\left[z \mid x\right] = \left(1 + \frac{\sigma_{xz}}{\sigma_x^2}\right) x.$$

If a researcher observes z, then they will notice that the forecast errors of the agent might be predicted by z:

$$\mathbb{E}\left[y - \hat{y} \mid z\right] \quad = \quad z\left(1 - \rho_{xz}^2\right),$$

where ρ_{xz} is the correlation between x and z. This expression makes it clear that the agent's forecast will be systematically related to realizations of z, as long as the correlation between x and z is not perfect. Intuitively, a higher correlation means that x contains more information about z, and thus the cost of not observing the latter decreases. At the same time, $\mathbb{E}\left[y - \hat{y} \mid x\right] = 0$ so the agent makes no systematic forecast errors conditional on their own observables.

2.2 Short Beliefs

Consider now y_t as following a stationary AR(2) process:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + v_t,$$

where v_t is i.i.d. white noise with strictly positive variance. Agents are interested in forecasting y_t, but they have short beliefs: they are constrained to use AR(1) models to form their forecasts. That is, they are not allowed to condition their forecasts on more than 1 lag, thus using a misspecified model.

As the number of observations goes to infinity, if they estimate the model $y_t = \psi y_{t-1} + \epsilon_t$ on the history $\{y_t\}_t$, the resulting coefficient would be:

$$\psi = \frac{\mathbb{E}\left[y_t y_{t-1}\right]}{\mathbb{E}\left[y_t^2\right]} = \frac{\phi_1}{1 - \phi_2}.$$

Let $\hat{y}_t := \psi y_{t-1}$ denote the forecasts produced with short beliefs. First note that this agent's forecast errors do not correlate with 1-period lags. That is,

$$\mathbb{E}\left[\left(y_t - \hat{y}_t\right) y_{t-1}\right] = \mathbb{E}\left[y_t y_{t-1}\right] - \left(\frac{\phi_1}{1 - \phi_2}\right) \mathbb{E}\left[y_t^2\right] = 0.$$

However, with 2-period lags:

$$\mathbb{E}\left[\left(y_t - \hat{y}_t\right) y_{t-2}\right] = \phi_2 \left[1 - \left(\frac{\phi_1}{1 - \phi_2}\right)^2\right] \mathbb{E}\left[y_t^2\right].$$

Since $\phi_1 \neq 1 - \phi_2$ is a necessary condition for stationarity, and $\mathbb{E}\left[y_t^2\right] > 0$, whenever $\phi_2 \neq 0$ short beliefs will lead agents to form forecast errors that are systematically correlated with 2-period lags.

2.3 Heterogeneous Persistence of Components

We will now consider one last example that will be helpful for the intuition in Section 4.3. Let y_t be a variable that an agent is interested in forecasting, which has two components that follow AR(1) processes:

$$\begin{aligned} y_t &= w_t + \varepsilon_t, \\ w_t &= \rho_w w_{t-1} + v_t^w, \\ \varepsilon_t &= \rho_\varepsilon \varepsilon_{t-1} + v_t^\varepsilon, \end{aligned}$$

where v_t^w and v_t^ε are both i.i.d. white noise, with strictly positive variances σ_w^2 and σ_ε^2 respectively. Consider an agent that only observes y_t and forecasts it with the AR(1) model $y_t = \rho_y y_{t-1} + v_t^y$. These are narrow beliefs in that the agent only observes a single variable instead of the entire vector of observables $\{y_t, w_t, \varepsilon_t\}$, and it turns out that this is relevant for predicting y_t. They are also short beliefs in that—whenever $\rho_w \neq \rho_\varepsilon$—it can be shown that y_t follows an ARMA(2,1) process, and so the agent is neglecting relevant lags when using an AR(1).

Again considering the asymptotic limit, the AR(1) coefficient that they will uncover will be:

$$\rho_y = \frac{\mathbb{E}\left[y_t y_{t-1}\right]}{\mathbb{E}\left[y_t^2\right]} = \frac{\left(\frac{\sigma_w^2}{1-\rho_w^2}\right)\rho_w + \left(\frac{\sigma_\varepsilon^2}{1-\rho_\varepsilon^2}\right)\rho_\varepsilon}{\frac{\sigma_w^2}{1-\rho_w^2} + \frac{\sigma_\varepsilon^2}{1-\rho_\varepsilon^2}}.$$

This expression has some intuitive properties. The persistence of y_t is a weighted average between the persistence of its two components. Moreover, the weights on the persistence of each component are given by their unconditional variance: we have $\sigma_w^2 / \left(1 - \rho_w^2\right) = \mathbb{E}\left[w_t^2\right]$ and the analogous expression for ε_t. So whatever component drives most of the variation in y_t will play a bigger role in shaping beliefs.

This is crucial because agents' forecasts will take the form $\hat{y}_{t+k} = \rho_y^k y_t$, so they might under- or overestimate the persistence depending on the shock at play. To see this, consider $\rho_w > \rho_\varepsilon$ and a starting point of $w_{t-1} = \varepsilon_{t-1} = 0$. If a shock $v_t^w > 0$ hits, then the full-information rational expectations conditional on all available information at time t is given by $\mathbb{E}_t\left[y_{t+k}\right] = \rho_w^k y_t$, and we know that $\rho_y \in (\rho_\varepsilon, \rho_w)$. Therefore, for any $k > 0$ we have $\hat{y}_{t+k} > \mathbb{E}_t\left[y_{t+k}\right]$. In this sense, expectations of future values of y_{t+k} overreact with shocks to w_t, and underreact with shocks to ε_t.[7] I will use this intuition to explain my

[7] The explanation I have provided is different to that of Coibion and Gorodnichenko (2015), nevertheless over- and underreaction in this context coincides exactly with their measure, which relates forecast errors to forecast updates.

results in Section 4.3.

Finally, a direct implication is that forecast errors will be predictable by the individual components:

$$\mathbb{E}\left[\left(y_{t+1}-\hat{y}_{t+1}\right)w_t\right]=\left(\rho_w-\rho_y\right)\frac{\sigma_w^2}{1-\rho_w^2},$$

and the analogous expression for ε_t shows the same kind of predictability.[8] However, one-period-ahead forecasts are not predictable by one lag of y_t, meaning $\mathbb{E}\left[\left(y_{t+1}-\hat{y}_{t+1}\right)y_t\right]=0$.

2.4 Forecast Errors and the Law of Iterated Expectations

In the three previous examples, an over-arching principle is that agents cannot detect any systematic errors in their forecasts when looking at them through the narrow and/or short lens of their models. Equivalently, their reduced information sets do not allow them to correct the systematic mistakes that are apparent to an outside observer with a larger information set.[9]

We can use this interpretation to connect to the literature attempting tests of rational expectations. Let $\hat{\mathbb{E}}\left[\cdot\right]=\mathbb{E}\left[\cdot\mid\Omega\right]$ denote the narrow and/or short expectations that arise when the information set is given by Ω. For an agent attempting to predict a variable y, if $x\in\Omega$, then the forecast errors of an agent who uses this information correctly have no systematic relationship with x:

$$\mathbb{E}\left[y-\hat{\mathbb{E}}\left[y\right]\mid x\right]=\mathbb{E}\left[y\mid x\right]-\mathbb{E}\left[\mathbb{E}\left[y\mid\Omega\right]\mid x\right]=0\quad\forall x,$$

where the last equality is due to the law of iterated expectations. The contrapositive of this statement is the key testable prediction produced by narrow and short beliefs: if $\mathbb{E}\left[y-\hat{\mathbb{E}}\left[y\right]\mid x\right]\neq 0$ for some x, then it must be that $x\notin\Omega$.

A simple way that the literature has found to test this is to estimate the moment

[8] Note that this property, as well as the overreaction, fails if $\rho_w=\rho_\varepsilon$. In such a case, the agents never make systematic mistakes because the persistence of the process is exactly what they expect: it's irrelevant whether the shock was on w_t or ε_t.

[9] This highlights two alternative interpretations of narrow and short beliefs. In the first one, agents are forecasting with misspecified models, but oblivious to their misspecification. In the literature, this is typically thought of as a deviation from rational expectations. However, in the alternative interpretation, agents have rational expectations but an information set that is truncated both the number of variables covered, and in the number of periods into the past. For the class of economies covered in this paper, both interpretations are equivalent and so both are welcome. However, the norm in the literature is to think about information sets as inevitably accumulating over time (which could be interpreted as no memory loss) but possibly incomplete in the number of variables. That is why I frame narrow beliefs as incomplete information, and short beliefs as an intertemporal misspecification. But this is just a matter of framing.

$\mathbb{E}\left[\left(y - \hat{\mathbb{E}}\left[y\right]\right) f(x)\right]$, by regressing the forecast error of y on $f(x)$, where $f(\cdot)$ is usually the identity function (but can be any functional form). This same approach can be used to discipline both narrow and short beliefs, if we have data on y and x as well as the expectations of agents: exclude variables and lags that, for some $f(\cdot)$, are found to be systematically related to forecast errors on key variables of interest. Much of the literature has already found these relations in one way or another, for example Kohlhas and Walther (2021) for expectations of output growth in the Survey of Professional Forecasters.

Importantly, the contrapositive mentioned above does not hold if an agent is forecasting with a parametric model that has the wrong functional form.[10] This implies that the strategy of excluding variables would not work for a parametric RPE, where agents could still make systematic forecast errors on a variable that they do include in their predictive models. This is the key relevance of the novel nonparametric generalization presented in this paper.

3 A General Framework

In order to introduce the equilibrium that results from narrow and short beliefs, I first outline a general environment for a class of models with heterogeneous agents that nests prominent examples in the literature, such as KS or heterogeneous-agent new Keynesian models. As an intermediate step, I define the full information and rational expectations (FIRE) equilibrium in this general framework to illustrate why it implies an infinite-dimensional state space for the agents in a general context. Then I consider how narrow and short beliefs shape the agent's decision problem, and finally I define the Nonparametric Restricted Perceptions Equilibrium (NRPE) that arises from narrow and short beliefs.

3.1 Environment

Time is discrete and infinite, denoted by t. There is a unit mass of ex-ante homogeneous agents[11] who are hit by idiosyncratic shocks ε_t, which follow a Markov process with transition kernel $\pi\left(\cdot \mid \varepsilon_t\right)$. They are ex-post heterogeneous in these shocks, as well as

[10] Consider an agent interested in predicting y which follows $y = x^2 + v$ where v is white noise. If the agent observes x and predicts y with the linear model $y = \varphi x + v$, then for some realizations of x the forecast will undershoot, and for others it will overshoot.

[11] This environment and the equilibrium definition can be easily generalized to include ex-ante heterogeneity, at the cost of additional notation. Since it is not the focus of the paper, I consider the ex-ante homogeneous case for exposition.

individual endogenous variables denoted by a_t, which have continuous support. At the same time, the economy is hit by aggregate shocks θ_t, which follow a Markov process with transition kernel $\Pi(\cdot \mid \theta_t)$. Aggregate equilibrium variables will thus be stochastic and denoted by x_t[12], and I will consider economies where they follow stationary and ergodic processes. I study atomistic agents in competitive markets, and thus they cannot individually affect the dynamics of x_t. But these aggregate variables are determined in general equilibrium and will be affected by the actions of all the agents.

At any given t, the agent's choice is next period's endogenous states a_{t+1}. These will determine the agent's utility given by $u(a_t, a_{t+1}, p_t)$ where p_t are the agents' directly-payoff-relevant variables, which they consider exogenously determined. They may be idiosyncratic or aggregate, and they can be written as a function of $\{\varepsilon_t, x_t, \theta_t\}$. I consider finite-dimensional p_t, and we can usually think of it as a low-dimensional object. For example, in the neoclassical economy described in Section 4.2, these would be the idiosyncratic efficiency units of labor and 2 aggregate prices: wages and the rental rate of capital.

The agents discount the future with factor $\beta \in (0, 1)$, and at time 0 they will try to maximize the expected discounted value of returns:

$$\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u(a_t, a_{t+1}, p_t) \right].$$

The operator $\mathbb{E}_t[\cdot]$ will be the focus of our analysis. It denotes subjective expectations, which coincide with objective expectations in the special case of FIRE. The agent's decision is constrained by the correspondence $a_{t+1} \in \Gamma(a_t, p_t)$ at every t. From this description alone we can see that only the expectations over p_t are directly relevant, expectations over other objects in $\{\varepsilon_t, x_t, \theta_t\}$ are only indirectly relevant if they help in forecasting p_t. Other equilibrium conditions necessary to complete the description of the environment (such as market clearing) will be discussed in conjunction with expectations below.

3.2 Full Information and Rational Expectations

If we want to study a general equilibrium, we need to pin down the expectations operator in order to compute optimal decisions. I will consider first the FIRE equilibrium notion and discuss it,[13] before studying narrow and short beliefs. Going forward, we

[12]This includes both aggregate states (determined at $t+1$) as well as forward-looking variables, and can even include moments of the distribution of individual states.

[13]For a treatment of these issues in a concrete and simple economy, KS is the canonical reference, and a similar economy is discussed in detail by Moll (2025)

focus on a recursive competitive equilibrium and so we drop time subscripts, using apostrophes to denote next period objects.

Let λ denote the joint distribution over a and ε. With full information and rational expectations in this class of economies, optimal decisions and market clearing lead to excess demand functions of the form:

$$ED\left(x, \lambda, \theta\right) = 0. \tag{1}$$

This condition can incorporate decisions of other agents such as representative firms or a government. It implies that generically, the directly-payoff-relevant aggregate variables p might depend on the entire distribution λ. Note that this latter object is infinite-dimensional, because of the continuous support over a.[14] For ease of exposition, I consider the case where this excess demand function implies that directly-payoff-relevant variables can be written as $p\left(\varepsilon, \lambda, \theta\right)$, and so without a dependence on x. This is exactly the case for the economy of Section 4 because the aggregate state can be obtained directly from λ, but it is easy to include the aforementioned dependence.[15]

A key characteristic of rational expectations is that agents understand this excess demand function and can use it to forecast p. With full information they additionally observe λ, as well as θ and x. Under these considerations, it follows that generically the recursive problem of the agents with value function $v\left(\cdot\right)$ follows:

$$\begin{aligned} v\left(a, \varepsilon; \lambda, \theta\right) = \max_{a'} & \left\{ u\left(a, a', p\left(\varepsilon, \lambda, \theta\right)\right) + \beta \mathbb{E}_{\pi, \Pi}\left[v\left(a', \varepsilon'; \lambda', \theta'\right) \mid \varepsilon, \theta\right] \right\} \\ \text{s.t.} \quad & a' \in \Gamma\left(a, p\left(\varepsilon, \lambda, \theta\right)\right), \\ & \lambda' = \Psi\left(\lambda, \theta\right). \end{aligned} \tag{2}$$

Where $\Psi\left(\cdot\right)$ is the law of motion for the distribution, which assumes that the distribution follows a joint Markov process with θ. But from the same recursive problem this is immediately verified, since given $\left(\lambda, \theta\right)$ we can compute the optimal actions that (in conjunction with π) pin down λ'. Another key assumption of rational expectations is that the agents perfectly know this law of motion, and use it to form expectations about λ'. But this is only indirectly necessary to forecast $p\left(\varepsilon, \lambda, \theta\right)$, and yet it implies that λ is

[14] A discrete support often leads to the same takeaways of this section in practice, because λ is high-dimensional in many applications of interest.

[15] This would be relevant if, for example, there were adjustment costs in investment. One would need to include a law of motion in the form $x'\left(x, \lambda, \theta\right)$ in the recursive problem, and x as an additional state variable.

a state variable in this problem. This is the key technical challenge that the literature has faced: (2) has an infinite-dimensional state, which standard recursive methods cannot handle.

Lastly, denote the agent's policy function by $a' = g(a, \varepsilon; \lambda, \theta)$. With these elements, we can define the FIRE equilibrium which will serve as a benchmark in the discussion.

Definition 1. *A FIRE Equilibrium is a set $\{v, g, p, \Psi\}$ such that:*

1. *Given $\{p, \Psi\}$, the policy g is optimal and v is the resulting value function in the Bellman equation* (2)*.*
2. *The variables p satisfy all equilibrium conditions embodied in* (1)*.*
3. *The law of motion Ψ is generated by $\{g, \pi\}$.*

Recall that condition 2 is a stand-in for many common equilibrium conditions: it can combine market clearing, optimality by static agents (such as representative firms), and government rules. The combination of conditions 1 and 3 implies the fixed point over Ψ that characterizes rational expectations: agents know the actual law of motion of the distribution, which is the relevant state space of the economy.

I finish this exposition with some discussion. In this class of models, two assumptions underlying FIRE seem particularly implausible. First is that agents—such as households—even observe infinite-dimensional objects in the first place, let alone pay attention to them. We can address this issue by considering an information structure that does not reveal infinite-dimensional objects to agents, and we would still be within the standard framework of rational expectations despite relaxing the full information assumption.[16] I call this approach *narrow beliefs*, as opposed to using the entire width of the available information at any point in time. As I discuss below, by itself that would still result in an infinite-dimensional state space because the infinite history of observations would replace the distribution as a state variable.

The second hardly-plausible assumption is that economic agents are able to process infinite-dimensional objects in a way that allows them to project their evolution into the future. Simply put, if trained economists have a hard time handling Ψ, it's difficult to believe actual households will. Much of the modern literature in bounded rationality instead argues that economic agents—even professional forecasters—use low-dimensional (misspecified) models to form their beliefs.[17] Formalizing this misspecification in the

[16] In fact, one of the seminal contributions of the rational expectations revolution—Lucas (1972)—pursues this approach.

[17] See Molavi (2024) for a discussion and a theory of simple models.

time dimension results in what I denote as *short beliefs*, and this is a deviation from rational expectations.

That said, while introspection might motivate the assumptions underlying narrow and short beliefs, the message of this paper is that we should go one step further and use empirical discipline to determine the extent of these frictions.[18] This approach, as covered in Section 2.4, is meant to guide the navigation through the "wilderness" outside FIRE. In some cases, it might be that equilibria in which we relax these implausible assumptions might look practically indistinguishable from the FIRE equilibrium, in which case we can always think of FIRE in an "as if" sense, as argued by Friedman (1953). We can interpret KS's finding of approximate aggregation in this light, I discuss this issue in more detail in Section 4.2.

Despite the strong assumptions pointed out, the FIRE equilibrium has many appealing properties for researchers. Most are aptly discussed by Sargent (2008) in his presidential address, but I will discuss some of them, which can be maintained while studying equilibria with narrow and short beliefs. First, beliefs are endogenously determined by the environment, which might include government policy rules, taxation, etc... and this grants robustness against the Lucas (1976) critique, so that we might be confident in using models to derive policy recommendations. Second, agents cannot be systematically fooled in equilibrium, which might be an appealing normative property. Third, subjective expectations are pinned down by the objective expectations given the economic environment, and thus the researcher has no degrees of freedom. This contrasts the adaptive expectations approach that preceded rational expectations, as argued in Sargent (1996). Fourth, FIRE yields sharp testable predictions not only for the equilibrium variables, but also for the agent's expectations. Fifth, in representative-agent models FIRE usually grants a great deal of analytical and computational tractability. This last property is no longer the case in heterogeneous-agent models. But for these reasons and others, FIRE remains an important benchmark to compare to when exploring theories of bounded rationality.

3.3 Narrow and Short Beliefs

Now I proceed to introduce two deviations from FIRE and define the NRPE for this class of economies. I briefly use time subscripts again to clarify the relevant intertemporal aspects of these deviations.

[18] A large literature has tested the FIRE hypothesis empirically and mostly rejected it, suggesting that narrow and short beliefs are promising. See Coibion et al. (2018) for a review in the context of inflation expectations.

Under narrow beliefs, agents do not observe $\{\varepsilon_t, \lambda_t, \theta_t\}$ every period, but they observe a finite set of variables $\hat{s}_t$ that are a function of $\{\varepsilon_t, x_t, \theta_t\}$,[19] such that their directly-payoff-relevant variables are perfectly revealed by $\hat{s}_t$.[20] Formally, a function p_s exists such that $p_t = p_s(\hat{s}_t)$ for any realization of $\hat{s}_t$. One example is provided by KS, who compute the equilibrium where agents observe their idiosyncratic labor efficiency units, aggregate productivity, and the first moment of the wealth distribution. In this case, the latter two objects can be mapped (via the firm optimality conditions) into wages and the rental rate of capital, which are directly relevant through the budget constraint. Another example would be agents simply observing p_t in the first place, in which case p_s is the identity function. As is evident, narrow beliefs are a constraint on the information set. I consider the case where a is still perfectly observed.

Under short beliefs, agents produce their forecasts using a finite number of lags of $\hat{s}_t$, as in Fuster et al. (2012). We will denote the number of lags as m, and consider the stacked formulation $s_t = \{\hat{s}_t, \hat{s}_{t-1}, ..., \hat{s}_{t-m+1}\} \in S$. We can then think of agents forecasting with a predictive model $Q : S \to \Delta(S)$, where the relevant uncertainty is over the realization of $\hat{s}_{t+1}$. The special case of $m = 1$ is denoted as a Markov model, and it is the case of KS. If the relevant number of lags is greater than m,[21] the agent is forecasting with a misspecified model and we abandon the realm of rational expectations. This is generically the case in this class of economies, where equilibrium objects such as prices do not follow Markov processes of any finite order, a point highlighted by Moll (2025).

Subject to these two constraints and dropping time subscripts again, we can write the recursive problem of the boundedly-rational agent as:

$$\begin{aligned} v(a,s) &= \max_{a'} \left\{ u\left(a, a', p(s)\right) + \beta \mathbb{E}_Q \left[v\left(a', s'\right) \mid s \right] \right\} \qquad (3) \\ \text{s.t.} \quad & a' \in \Gamma\left(a, p(s)\right), \\ & s' \sim Q(s). \end{aligned}$$

In this formulation, the $\mathbb{E}_Q[\cdot \mid s]$ operator is meant to denote expectations generated by the model Q, given s which contains m lags of $\hat{s}$.[22] Q fulfills the role of a perceived law of motion, in particular over variables and lags that are allowed by the narrow and short

[19]Recall that x_t are all of the aggregate equilibrium variables, which can include moments of the distribution.

[20]This rules out imperfect perception of prices such as in Gabaix (2014).

[21]Meaning that including more lags would increase the predictive accuracy.

[22]It is easy to extend this formulation to have a different number of lags for different variables, which I avoid for the sake of exposition. This can also be viewed as a restriction on Q, and so the current approach is without loss.

beliefs of the agent. It can be interpreted as the agent's (misspecified) predictive model.

It is immediately noticeable that the agent's state space is finite-dimensional: $\hat{s}$ is a finite set, and s includes a finite number of lags of it. Therefore, this problem can be solved with standard recursive methods, given Q. Importantly, both frictions are necessary for this result: if the agent had full information then the infinite-dimensional distribution would be a relevant state. But even if the agent has narrow beliefs, the infinite-dimensional history of equilibrium objects that they observe could still a relevant state, because they are not Markov. Another way to see this point is that the history of equilibrium variables could be informative about the shape of the distribution. I show one example numerically in Appendix A.2. Therefore short beliefs are also necessary for the computational payoff.

We will pin down Q through the equilibrium definition, but we need some preliminaries. If we denote the agent's resulting policy function by $a' = g(a, s)$, as before that will imply a law of motion for the distribution Ψ. In this equilibrium it will also be the case that the state space of the economy is $\{\lambda, \theta\}$, which will now be different from the agent's reduced state space. The implication is that Ψ and Π govern the evolution of the economy's state, and so they will pin down the ergodic distribution of all equilibrium variables. Let $\mu_s \in \Delta(S)$ denote the ergodic distribution of s, and let $\mu \in \Delta(S^2)$ denote the joint distribution of s and s', where the marginal distribution of each of those elements is the ergodic distribution μ_s. That is, $\mu_s(\mathcal{S} \times \mathcal{S}')$ tells us the long-run probability of observing $s \in \mathcal{S}$ and next period $s' \in \mathcal{S}'$. With some abuse of notation, let $Q(\mathcal{S}' \mid s) := \mathbb{P}_Q[s' \in \mathcal{S}' \mid s]$, where $\mathbb{P}_Q$ denotes the probability under Q. We are now ready to define the equilibrium with narrow and short beliefs.

Definition 2. *A Nonparametric Restricted Perceptions Equilibrium (NRPE) is a set* $\{v, g, p, \Psi, \mu, Q\}$ *such that:*

1. *Given Q, the policy g is optimal and v is the resulting value function in the Bellman equation* (3).
2. *The variables s and $p(s)$ satisfy all equilibrium conditions embodied in* (1).
3. *The law of motion Ψ is generated by $\{g, \pi\}$.*
4. *The distribution μ is induced by $\{\Psi, \Pi\}$.*
5. *For any measurable $\mathcal{S}, \mathcal{S}'$ in the Borel sigma-algebra of S:*

$$\int_{\mathcal{S}} Q(\mathcal{S}' \mid s)\, \mu_s(ds) = \mu(\mathcal{S} \times \mathcal{S}'), \tag{4}$$

where $\mu_s(\mathcal{S}) = \mu(\mathcal{S}, S)$.

Equation (4) is a belief consistency condition: the conditional probabilities over next period's observables given current observables must be consistent with the realized distribution of those variables in equilibrium. In the language of KS, the perceived law of motion must be consistent with the actual law of motion. In fact, these authors define a FIRE equilibrium, but compute an RPE where Q takes on a particular parametric form, that happens to approximate FIRE extremely well. I define this RPE in Appendix A.1.

A couple of issues now merit discussion. Analogously to the FIRE equilibrium where we looked for a fixed point over the law of motion for the distribution, here the fixed point is over the law of motion for the agent's observables and follows the same logic: beliefs influence actions, which determine equilibrium outcomes, which pin down beliefs. It is also the case that RPE in general consist of augmenting a temporary equilibrium (in the sense of Grandmont, 1977) with expectations that align with endogenous outcomes of the equilibrium. This grants them the same kind of robustness to the Lucas critique as FIRE.[23] And given the frictions, beliefs are pinned down by the environment, providing testable predictions. So are narrow and short beliefs a free parameter, that researchers must beware?[24] This is not the case as long as we use data on expectations to bring empirical discipline to this equilibrium notion, as argued in Section 2.4.

In the NRPE beliefs are fully internally consistent, meaning that the law of iterated expectation holds, and agents don't make any systematic forecast errors on s' conditional on s. Their forecast errors may be predicted by variables that they do not observe, or longer lags than what their perceived law of motion allows, but they can't do anything about it under their current constraints. I explore this implication in the next section.

To the best of my knowledge, this paper is the first to study a nonparametric generalization of an RPE, and it has some interesting implications. First, it removes a degree of freedom to the researcher whenever the perceived law of motion does not have a perfect fit to the realized law of motion, because in that case the choice of parametric form matters for the equilibrium outcomes. In the case of KS, this is solved by selecting the one that most closely approximates FIRE, which is the goal, unlike this paper and others in the bounded rationality literature. Second, it allows for the agents to form accurate expectations (given their constraints) over their variables of interest when they follow complicated processes. These may be poorly approximated by standard paramet-

[23] An endogenous reaction that might be missing from this framework is information acquisition, as in Broer et al. (2021), since policy might change information sets. I argue in the conclusion that this is a promising avenue for future research on narrow and short beliefs.

[24] This language echoes the point made in the introduction of Hansen and Sargent (2008)

ric forms, which might be the case with many equilibrium variables.[25]

As a third implication, an NRPE poses no tension between statistical and normative mistakes, as opposed to an RPE where the parametric form does not provide an exact fit to the equilibrium law of motion. In this case, there will be some systematic forecasting mistakes which will be minimized following some measure of statistical fit, like ordinary least squares. Generically, these least-square errors are not equal to the utility mistakes made by agents forecasting with misspecified models, and so they are not the welfare-maximizing parameters.[26] In a FIRE equilibrium this tension is not present because all forecast errors are purely random, driven by the environment's inherent uncertainty. Similarly, in an NRPE all forecast errors are purely random conditional on s, and so the absence of a parametric form erases the systematic mistakes that create the tension between statistical fit and welfare.

Lastly, proving existence and uniqueness of the NRPE in heterogeneous-agent economies is beyond the scope of this paper, as this has proven to be a tremendously difficult endeavor for even the relatively simple economy of KS. See Cao (2020) for recent results and discussion of the technical issues, while still not achieving a proof of the existence of the recursive equilibrium with the natural minimal state space of $\{\lambda, \theta\}$. For the parametric RPE in representative-agent models, proof techniques exist at least since Marcet and Sargent (1989), and are well covered by Evans and Honkapohja (2012).

3.3.1 Computational Algorithm

In order to find the computational solution to this equilibrium, one may use an algorithm similar to that of KS, since the relevant fixed point is over the perceived and realized law of motion for aggregate variables. However, there are nuances in the actual implementation, which merit substantial innovation. First of all, how to implement the nonparametric predictive model Q computationally? I propose to use a Markov chain of order m, with as many gridpoints as needed to ensure computational accuracy. In particular, one may start by guessing grids $\boldsymbol{s}_0$ and transition probabilities $\boldsymbol{Q}_0$. Then, for iterations $j \geq 0$:

1. Given $\left\{\boldsymbol{s}_j, \boldsymbol{Q}_j\right\}$, solve the recursive problem (3) to get the optimal policy g
2. Use g as well as any other equilibrium conditions to simulate the economy for

[25] For example, in models with crises driven by multiple equilibria (such as Gertler et al., 2020) equilibrium variables may exhibit multimodality. See Adrian et al. (2021) for a discussion.

[26] Estimating welfare-maximizing parameters would usually be a daunting computational task, in contrast many of our statistical methods rely on measures of fit that are amenable to cheap computation.

many periods.[27]

3. Approximate the dynamics of the simulated equilibrium history $\{s_t\}_t$ with a Markov chain $\left\{\boldsymbol{s}_{j+1}, \boldsymbol{Q}_{j+1}\right\}$

4. If $\boldsymbol{s}_{j+1} \approx \boldsymbol{s}_j$ and $\boldsymbol{Q}_{j+1} \approx \boldsymbol{Q}_j$, stop. Otherwise, go back to step 1 for $j+1$.[28]

Note that—as long as it is excluded from the information set of the agent—the aggregate shocks θ might be sampled from a continuous distribution. This is in contrast to many FIRE equilibria and the KS algorithm, in which the agent's state space must be finite to use standard recursive computational methods. In addition, θ might be relatively high-dimensional (as in medium to large-scale Dynamic Stochastic General Equilibrium (DSGE) models), but as long as s has relatively low dimension, our recursive methods will not be subject to the costly curse of dimensionality. While I do not exploit this advantage in this paper, it is a promising feature of the NRPE for future work.

Step 3 leaves open the question of how exactly to approximate the simulated data with a Markov chain. Most of the numerical methods in the literature consider approximations of parametric processes, like the linear Gaussian case covered by Tauchen (1986) or Rouwenhorst (1995). This does not serve this paper's nonparametric needs, so I propose a novel histogram-based approximation method that can be applied to any multivariate stochastic process from which we can draw samples, such as our equilibrium objects s_t.[29] It is simple, highly parallelizable, and successfully matches persistences that are close to unit root. I explain this approach in Appendix C.

4 Application: Narrow Beliefs and Consumption Dynamics

In this section, I outline a neoclassical model with heterogeneous households who face uninsurable idiosyncratic income shocks, similar to the economy in KS. I focus on the implications of narrow beliefs over income components as in Section 2.3, by comparing the aggregate dynamics to a benchmark equilibrium that approximates FIRE.

Toward this goal, I first show that if households observe their idiosyncratic efficiency units of labor in addition to wages and the rental rate of capital, narrow and short beliefs

[27] Burn-in a number of initial periods.

[28] It may be desirable to dampen the update for stability purposes.

[29] To the best of my knowledge, there are two options in the literature: Farmer and Toda (2017), and Janssens and McCrary (2023). In future work, I plan to explore the differences between my method and theirs.

have no economic bite, and the NRPE is very similar to the FIRE equilibrium. I explore this similarity theoretically by showing that this NRPE is equivalent to a generalization of the KS RPE—which we know approximates FIRE very well. Therefore, I interpret this NRPE as a very close approximation to FIRE, and I use it as a benchmark.

Then, I introduce narrow(er) beliefs over income components: agents cannot tell apart their efficiency units of labor from wages. This friction actually has bite and will lead to an NRPE where aggregate consumption underreacts to aggregate productivity shocks, relative to the FIRE approximation. This mechanism allows for a better match to the dynamics of consumption and investment that we observe in the data at a business cycle frequency.

4.1 Environment

This economy is a simple special case of the class of models studied in Section 3, and similar to KS. A unit mass of ex-ante homogeneous households with constant relative risk aversion of $\gamma > 0$ and discount factor $\beta \in (0,1)$ seek to maximize the expected discounted utility of consumption. They earn labor income by supplying efficiency units of labor ε, which are an idiosyncratic, stochastic and uninsurable Markov process with transition kernel $\pi(\cdot \mid \varepsilon)$. Labor is paid a wage of w per efficiency unit. In addition, households may save in claims to capital denoted by a, with a short-selling constraint of $\underline{a} > 0$. Capital depreciates at rate $\delta \in (0,1)$ and is paid a rental rate of r. Letting a' denote next period's capital holdings, the constraints faced by households are given by:

$$\begin{aligned} c + a' &= w\varepsilon + (1 + r - \delta)\, a, \\ a' &\geq -\underline{a}. \end{aligned} \tag{5}$$

A representative competitive firm produces the final good that is used for consumption or future capital. Aggregate output Y follows from the Cobb-Douglas production function:

$$Y = \theta K^{\alpha} H^{1-\alpha},$$

where $\alpha \in (0,1)$ is the capital share of output. θ is an aggregate productivity shock, that evolves with transition kernel $\Pi(\cdot \mid \theta)$.

Firm Optimality and Market Clearing Since firms are representative and competitive, their input prices simply equal their marginal product:

$$\begin{aligned} w &= \theta(1-\alpha)\left(\frac{K}{H}\right)^{\alpha}, \\ r &= \theta\alpha\left(\frac{K}{H}\right)^{\alpha-1}. \end{aligned} \tag{6}$$

Let $\lambda(a,\varepsilon)$ denote the distribution over capital holdings and efficiency units of labor. In equilibrium, labor, capital and goods markets clear, respectively:

$$\begin{aligned} H &= \int \varepsilon\lambda(d\varepsilon), \\ K &= \int a\lambda(da), \\ Y+(1-\delta)K &= \int \left[c(a,\varepsilon)+a'(a,\varepsilon)\right]\lambda(da,d\varepsilon). \end{aligned} \tag{7}$$

These conditions will always be satisfied in the equilibria we consider. Households will also optimize, what remains to be determined is their expectations. For completeness, aggregate consumption is $C := \int c(a,\varepsilon)\lambda(da,d\varepsilon)$ and investment is $I := K' - K(1-\delta)$.

4.2 The Benchmark Equilibrium: a FIRE-Equivalence Result

We shall first consider the case where the households observe $\{\varepsilon, w, r\}$ every period. They have narrow beliefs because they do not observe λ, but in the next section they will be narrower. In addition, we consider short beliefs of 1 lag: the household will only use $\{\varepsilon, w, r\}$ to forecast $\{\varepsilon', w', r'\}$. So in the framework of Section 3.3, $s = \{\varepsilon, w, r\}$.

It is immediate that in the NRPE, the households will effectively know π because ε follows a Markov process. Then, we can formulate the household's recursive problem as:

$$\begin{aligned} &v(a,\varepsilon;w,r) = \max_{c,a'}\left\{\frac{c^{1-\gamma}}{1-\gamma} + \beta\mathbb{E}_{\pi,Q}\left[v\left(a',\varepsilon';w',r'\right) \mid \varepsilon;w,r\right]\right\} \\ &\text{s.t.} \quad \left\{w',r'\right\} \sim Q(w,r), \end{aligned} \tag{8}$$

and subject to the constraints (5). That is, households observe current prices, and use them to form expectations over future prices with their predictive model Q. Since π is fully revealed, we have reserved Q for the prices only. This is without loss and still fits the framework of Section 3.3, but this alternative formulation will ease the exposition.

Going forward, we denote the prices as $p = \{w, r\}$. Note that excluding ε, they are the only directly-payoff-relevant variables here.

For $p \in P$, let $\mu \in \Delta\left(P^2\right)$ denote the joint ergodic distribution of p and p', as in the general framework. This allows us to define the benchmark equilibrium as follows:

Definition 3. *The benchmark Nonparametric Restricted Perceptions Equilibrium (NRPE) is a set $\{v, g, p, \Psi, \mu, Q\}$ such that:*

1. *Given Q, the policy g is optimal and v is the resulting value function in the Bellman equation* (8)*.*
2. *Prices p satisfy firm optimality* (6)*.*
3. *Labor, capital and goods markets clear, satisfying* (7)*.*
4. *The law of motion Ψ is generated by $\{g, \pi\}$.*
5. *The distribution μ is induced by $\{\Psi, \Pi\}$.*
6. *For any measurable $\mathcal{P}, \mathcal{P}'$ in the Borel sigma-algebra of P:*

$$\int_{\mathcal{P}} Q\left(\mathcal{P}' \mid p\right) \mu_p\left(dp\right) = \mu\left(\mathcal{P} \times \mathcal{P}'\right), \tag{9}$$

where $\mu_p\left(\mathcal{P}\right) = \mu\left(\mathcal{P}, P\right)$.

How severe is the bounded rationality of the agents in this equilibrium? One may be worried since they can only observe two prices instead of the entire distribution of wealth and aggregate productivity, which correspond to the natural minimal state space in FIRE for this economy, as in KS. However, it turns out that this information set and lag length allows the agents to forecast very well. The reason is that the firm optimality conditions provide an invertible mapping between $\{w, r\}$ and $\{K, \theta\}$, which implies that these two sets of variables are informationally equivalent in prediction problems. And, we know from KS that using $\{K, \theta\}$ to forecast equilibrium prices works extremely well, therefore this result extends to our benchmark equilibrium as well. It implies that we can think of this benchmark equilibrium as a very good approximation to FIRE. This is the interpretation that I hold moving forward.

This can be formalized by considering a generalization of the RPE that KS compute. Their original RPE (which I define and discuss in Appendix A.1) features a perceived law of motion of the form:

$$K' = f_\phi\left(K, \theta\right),$$

where $f_{\phi}(\cdot)$ is a parametric function, with parameters ϕ. Note that this produces a point forecast over capital next period. This and the parametric form can be relaxed by considering a perceived law of motion of the form $Q_{KS} : \boldsymbol{K} \times \Theta \to \Delta(\boldsymbol{K})$ for $K \in \boldsymbol{K}$ and $\theta \in \Theta$. In a KSNRPE equilibrium (which I define formally in Appendix B) agents solve their recursive problem with the perceived law of motion Q_{KS}, and Q_{KS} satisfies a belief consistency condition analogous to (9). I include the proof to the following equivalence result.

Proposition 1. *For any benchmark NRPE, there exists a KSNRPE with the same* Ψ*, and vice-versa.*

The proof is given in Appendix B. In this environment, having the same Ψ means that both equilibria are completely observationally equivalent to a researcher who is observing realizations of equilibrium variables. That is because the state space of the economy is still $\{\lambda, \theta\}$, and so with the same Ψ the state spaces evolve with identical dynamics, producing the same equilibrium variables, and pinning down informationally equivalent predictive models for the agents.

Note that in a KSNRPE, agents can forecast weakly better than in the original KS equilibrium, because the perceived law of motion is allowed more flexibility in a way that cannot hurt the agent. Therefore, when we consider the benchmark NRPE, we are approximating FIRE even better than KS did, who already did a fantastic job.[30] Consequently, henceforth I interpret the equilibrium of definition 3 as the FIRE benchmark, against which to interpret narrower beliefs.

Lastly, this result can be viewed as a microfoundation for the equilibrium computed by KS, in that it might be more plausible to think that households observe prices instead of aggregate productivity or moments of the distribution. Because of the equivalence, we can always think in "as if" terms when considering the KS equilibrium.[31]

4.3 Narrow Beliefs over Income Components

This section now considers the implications of even narrower beliefs. I take as given an empirical finding by Adams and Rojas (2024): households cannot differentiate whether their income changes were driven by idiosyncratic or aggregate sources. These authors

[30]In this sense, their constraints on a point forecast and particular parametric form are mostly irrelevant in quantitative terms. I consider the generalization purely for theoretical purposes.

[31]This section's result and takeaway pertain only to the particular environment under study. One can easily think of economies in which prices do not reveal the relevant aggregate states. A simple example would be if θ followed an AR(2) process, and so the relevant state space of the economy would include a lag of θ which would not be uncovered by observing $\{w, r\}$.

study the persistence of each component, and the expected persistence that is implied by their income expectations, reported to the Survey of Consumer Expectations. They find that households do not adjust their expectations in the manner predicted by FIRE: they do not correctly anticipate the persistence of the shock that is hitting them.[32]

The NRPE allows us to study this phenomenon very naturally in a general equilibrium model of a closed economy.[33] In the current environment $w\varepsilon$ is total labor income in the budget constraint, where w is aggregate and ε is idiosyncratic. Therefore, we can study households that observe $y = w\varepsilon$ instead of w and ε separately. Under the framework of Section 3.3 we have $p = s = \{y, r\}$, so the agent observes less variables than in the FIRE approximation that we just covered. For that reason, the equilibrium in this section will henceforth be referred to as "narrow beliefs".[34] Under these beliefs, the recursive problem of the household can be framed as follows.

$$
\begin{aligned}
v(a; y, r) &= \max_{c,a'} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_Q \left[v\left(a'; y', r'\right) \mid y, r \right] \right\} \qquad (10) \\
\text{s.t.} \quad & \left\{y', r'\right\} \sim Q(y, r), \\
& c + a' = y + (1 + r - \delta) a, \\
& a' \geq -\underline{a}.
\end{aligned}
$$

Why might narrow beliefs matter for the agent's decision? We can go back to the intuition from Section 2.3: if agents don't observe income components, then they will expect their income y's persistence to be a weighted average between the persistences of w and ε. In particular, we know that most of the volatility in household-level income comes from idiosyncratic sources, as opposed to aggregate,[35] so the persistence of ε will have a stronger influence on household's expectations over y' given y. Consider w as more persistent than ε, then whenever w increases because of an aggregate shock, all households will underestimate the persistence of the increase. This is because they only observe their y increasing, and it usually reverts to the mean with the persistence of ε. Consequently, they will underestimate the increase in the present value of their future income relative to the case of full information. A simple permanent income logic tells

[32]While these authors do not test the predictability of forecast errors, this is a direct consequence of the tests they run.

[33]Adams and Rojas (2024) study it in a small open economy, a simpler endeavor due to exogenous prices. I discuss how my results compare to theirs in Section 4.5.

[34]As discussed in Section 2.3, the short beliefs friction bites too. Labeling this equilibrium as narrow beliefs only stresses the fact that relative to FIRE, the agent has access to a lower number of observables, for the same lags.

[35]This will be the result of my calibration, but it has been extensively documented empirically, for example in Pischke (1995) in the same context of narrow beliefs.

us that household consumption will underreact. Since this is the case for all households, aggregate consumption will follow the same dynamics.

I close this section by noting that the equilibrium with narrow beliefs follows definition 3, but with the Bellman equation given by (10) and $p = \{y, r\}$.

4.4 Calibration and Solution

I calibrate the model at a quarterly frequency and choose parameter values that are standard in the literature. I set the discount factor at $\beta = 0.99$, the relative risk aversion at $\gamma = 4$,[36] the capital share at $\alpha = 1/3$, the depreciation rate at $\delta = 0.025$, and the short-selling constraint at $\bar{a} = 0$.

For the difference between the FIRE benchmark and the narrow beliefs economy, the most crucial parameters are those governing the aggregate shocks θ and idiosyncratic shocks ε. I assume both of these follow AR(1) processes in logs, and in the Fernald (2012) data I find an autoregressive coefficient of $\rho_\theta = 0.99$ and volatility of $\sigma_\theta = 0.008$ for the deviations from a log-linear trend.

For the labor efficiency units, there is more uncertainty regarding the persistence. While some of the literature finds that persistent components of income follow a unit root,[37] others find much lower estimates such as Guvenen (2009) who estimate a heterogeneous income profile process with persistent and transitory components. After subtracting the income profiles, the estimated persistence of the process is given by 0.856 at a quarterly frequency.[38] For the main exercise I consider $\rho_\varepsilon = 0.95$ which is slightly above the midpoint of the range of estimates. I then set $\sigma_e = 0.236$ to match the variance of 5-year income growth rates, as documented by Guvenen et al. (2014).

I solve the model globally for the FIRE benchmark equilibrium, and the narrow beliefs equilibrium, with the algorithm of Section 3.3.1. I use the endogenous gridpoint method of Carroll (2006) to solve the recursive problem of the households, and the Young (2010) method to simulate the economy. I draw θ from its continuous distribution

[36] This is common in the literature with heterogeneous agents and aggregate shocks, see Bayer et al. (2024) and Broer et al. (2021).

[37] An early influential example was Abowd and Card (1989)

[38] His residual process is given by a permanent and a transitory component: $\varepsilon_t = \varepsilon_t^P + \varepsilon_t^T$, where $\varepsilon_t^P = \rho_P \varepsilon_{t-1}^P + v_t$ and both ε_t^T and v_t are i.i.d. white noise with variance σ_T^2 and σ_P^2 respectively. Since I do not have a transitory component, the relevant statistic is the first-order persistence of ε_t, given by:

$$\frac{\mathbb{E}\left[\varepsilon_t \varepsilon_{t-1}\right]}{\mathbb{E}\left[\varepsilon_t^2\right]} = \frac{\rho_P \sigma_P^2}{\sigma_P^2 + \left(1 - \rho_P^2\right) \sigma_T^2}.$$

I calculate this value, given his reported estimates of $\rho_P, \sigma_P^2, \sigma_T^2$

in this simulation, which is computationally feasible since it is not part of the household's state space, as opposed to the usual case with the KS method. The stochastic steady state (SSS) of the economy is found by simulating the economy in the absence of any aggregate shocks, until the distribution λ converges.

4.5 Results

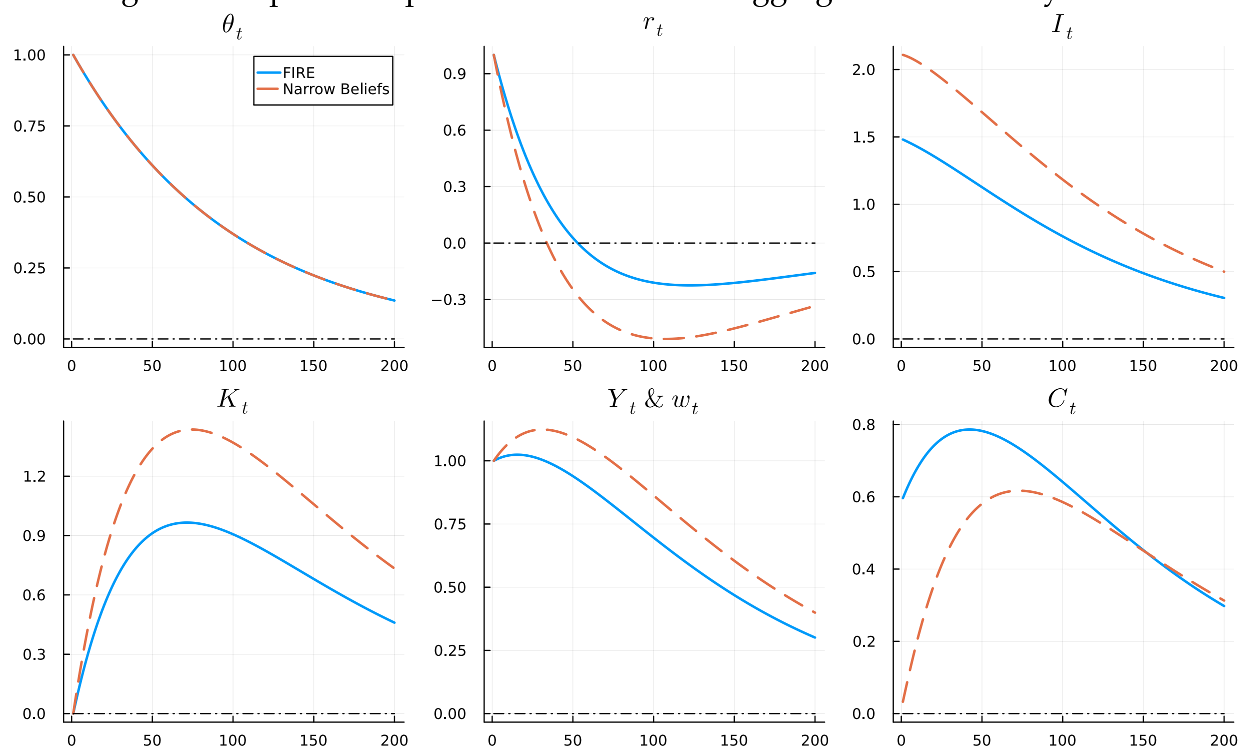
The main implication of narrow beliefs comes in the aggregate dynamics following a productivity shock. I illustrate the mechanism by computing the impulse response functions of a 1% shock to θ starting from the SSS, shown in figure 1. A shock to θ raises household income through wages,[39], which raises consumption and investment. However, in the narrow beliefs economy, households attribute most of the extra income to an idiosyncratic shock, leading them to underestimate its persistence—and consequently their permanent income. This leads to a lower consumption response on impact, relative to FIRE, and a higher investment response, which triggers capital accumulation. Higher K further raises wages in the future, leading to persistent underreaction of consumption and overreaction of investment, even after most of the θ shock has faded away. This implies that the persistence of the aggregate component of income of the household is greater than the persistence of the underlying shock, providing amplification to the strength of this mechanism. In particular, it is notable how investment overreacts persistently despite a large decrease in the returns to investment as evidenced by the reaction of r, which the agents observe.

At this point, the reader might think that it's implausible that households would be systematically fooled as in this impulse response, where they are overly pessimistic in that they underestimate the persistence of their income gains for many consecutive periods. Two things are worth pointing out. First is that when the economy is continuously hit by shocks, agents will sometimes be overly optimistic as well, in thinking that negative shocks will die out quickly. Second, when idiosyncratic shocks hit them (and these drive most of their income volatility), they will always overestimate the persistence of their income gains, leading to overly optimistic responses to income gains and pessimistic responses to income losses, contrary to the effect of aggregate shocks. On average, agents are not systematically mistaken when conditioning only on their total income, these mistakes are only evident when conditioning on the source of the change.

After understanding the different dynamics of consumption and investment predicted by narrow beliefs, we want to compare these to the observed aggregate data.

[39]When measured as % deviations from the SSS, the response of output always coincides with that of wages in this environment. This occurs because they share the elasticity to capital α, and H is constant.

Figure 1: Impulse Response Functions to an Aggregate Productivity Shock



All variables expressed in percentage deviation from their stochastic steady state values. The horizontal axis denotes time periods, the model is solved at a quarterly frequency.

Table 1: Consumption and Investment Moments

	$\sigma(C)/\sigma(I)$	$\sigma(\Delta C)/\sigma(\Delta I)$	$cor(C,I)$	$cor(\Delta C, \Delta I)$
FIRE	0.36	0.35	0.94	0.93
Narrow Beliefs	0.2	0.21	0.43	0.35
Data	0.19	0.22	0.65	0.29

In the following exercise I include output in the statistics for completeness, but the simple environment under study does not feature many of the known drivers of economic activity that are present in larger dynamic stochastic general equilibrium models,[40] and thus I stress the relative dynamics between consumption and investment as the main feature that I study. I obtain these series in US data from FRED for the period 1947:Q1 to 2025:Q2,[41] and I employ an HP-filter on their natural logs at a frequency of 1,600 to study the movements at a business cycle frequency, as is standard in the literature. I use this same filter on model-simulated data to compare, so all moments measure units in log deviations from the HP trend.

Table 1 shows the main results of this exercise. While in the FIRE equilibrium consumption is too volatile relative to investment, the narrow beliefs equilibrium hits this moment very well. Remarkably, this is true both when computing the volatilities of the levels as well as the volatilities of the innovations, where $\Delta X_t := X_t - X_{t-1}$. FIRE also produces too much correlation between consumption and investment. Narrow beliefs matches this moment quite well for the innovations, and for the levels it "overshoots" suggesting a correlation that is too low relative to the data, but still closer to it than FIRE. The main takeaway is that the joint distribution of consumption and investment seems to be better matched by narrow beliefs.

Appendix D shows that the rest of the business cycle moments—in levels as well as innovations—are broadly similar across the equilibria, except that in FIRE consumption is overly procyclical while in the narrow beliefs it is not sufficiently so, by roughly the same amount. These results expand on the literature that has studied narrow beliefs over income components in settings where labor earnings and the return on savings are determined exogenously. Pischke (1995) first studied this for consumers with quadratic utility and no borrowing constraints in an endowment economy, which Ludvigson and Michaelides (2001) expanded to include buffer-stock savers, and both papers find that

[40] From an accounting perspective, this environment does not have government expenditures or an external sector. From an economic perspective, there are no aggregate demand shocks which might be relevant for business cycle volatility.

[41] In particular, I use the codes GDPC1, GPDIC1, and PCECC96 for output, investment and consumption respectively

narrow beliefs over income components are insufficient to explain the degree of consumption smoothness that we observe in aggregate data, whereas it seems to be sufficient in this general equilibrium setup that features an amplification mechanism. On the other hand, Adams and Rojas (2024) argue that the idiosyncratic component of income is more persistent, and use narrow beliefs to generate excess consumption volatility in small open economies.[42] The data that this paper looks at requires smoother consumption, which is delivered by the calibration that I use.

5 Conclusion

This paper studies a novel equilibrium notion that arises from considering two empirically-supported deviations from full information and rational expectations. In models with heterogeneous-agents and aggregate uncertainty, narrow and short beliefs result in a finite-dimensional state space for the agents in the economy, unlike FIRE. Moreover, we can bring empirical discipline to these frictions by analyzing the predictability of forecast errors.

The computational bite of these frictions (on the dimension of the state space) may be very different from the economic bite. In my application in a neoclassical model with heterogeneous-agents, I show first that the Krusell and Smith (1998) equilibrium is equivalent to one where agents observe prices and forecast with a single lag, and so both are very close to FIRE. However, when I incorporate recent evidence on narrow beliefs over income components, I show large differences in the dynamics of consumption and investment that better match the data.

I am studying further issues in ongoing work. First, I am analyzing the predictability of forecast errors when conditioning on income components, by using the approach of Section 2.4 to study the application. Second, I am studying the implications of narrow beliefs for further issues, such as the cost of business cycles. Third, I am exploring richer features in the neoclassical model, such as an endogenous labor supply margin. The last issue is a computational challenge for the literature, but the NRPE shows promise in being amenable to efficient computation in this regard.

[42]This relies on the inverse logic of this paper: when an aggregate shock hits, agents overestimate its persistence and consumption overreacts.

References

Abowd, J. M. and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57, 411–445, publisher: [Wiley, Econometric Society].

Adams, J. J. and E. Rojas (2024): "Household Consumption and Dispersed Information," *Journal of Monetary Economics*, 147, 103592.

Adrian, T., N. Boyarchenko, and D. Giannone (2021): "Multimodality in Macrofinancial Dynamics," *International Economic Review*, 62, 861–886, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12501.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2021): "Imperfect Macroeconomic Expectations: Evidence and Theory," *NBER Macroeconomics Annual*, 35, 1–86, publisher: The University of Chicago Press.

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021): "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, 89, 2375–2408, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA17434.

Azinovic, M., L. Gaegauf, and S. Scheidegger (2022): "Deep Equilibrium Nets," *International Economic Review*, 63, 1471–1525, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/iere.12575.

Baley, I. and J. Turen (2023): "Lumpy Forecasts," .

Bayer, C., B. Born, and R. Luetticke (2024): "Shocks, Frictions, and Inequality in US Business Cycles," *American Economic Review*, 114, 1211–1247.

Boppart, T., P. Krusell, and K. Mitman (2018): "Exploiting MIT shocks in heterogeneous-agent economies: the impulse response as a numerical derivative," *Journal of Economic Dynamics and Control*, 89, 68–92.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–2782.

Bordalo, P., N. Gennaioli, and A. Shleifer (2022): "Overreaction and Diagnostic Expectations in Macroeconomics," *Journal of Economic Perspectives*, 36, 223–244.

Branch, W. A. and G. W. Evans (2010): "Asset Return Dynamics and Learning," *The Review of Financial Studies*, 23, 1651–1680.

Broer, T., A. Kohlhas, K. Mitman, and K. Schlafmann (2021): "Information and Wealth Heterogeneity in the Macroeconomy," .

Cao, D. (2020): "Recursive equilibrium in Krusell and Smith (1998)," *Journal of Economic Theory*, 186, 104978.

Carroll, C. D. (2006): "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 91, 312–320.

Coibion, O. and Y. Gorodnichenko (2015): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 105, 2644–2678.

Coibion, O., Y. Gorodnichenko, and R. Kamdar (2018): "The Formation of Expectations, Inflation, and the Phillips Curve," *Journal of Economic Literature*, 56, 1447–1491.

Den Haan, W. J., K. L. Judd, and M. Juillard (2010): "Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty," *Journal of Economic Dynamics and Control*, 34, 1–3.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2021): "Exposure to Grocery Prices and Inflation Expectations," *Journal of Political Economy*, 129, 1615–1639, publisher: The University of Chicago Press.

Evans, G. W. and S. Honkapohja (2012): *Learning and Expectations in Macroeconomics*, Princeton University Press.

Farmer, L. E. and A. A. Toda (2017): "Discretizing nonlinear, non-Gaussian Markov processes with exact conditional moments," *Quantitative Economics*, 8, 651–683, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/QE737.

Fernald, J. G. (2012): "A quarterly, utilization-adjusted series on total factor productivity," *Working Paper Series*, number: 2012-19 Publisher: Federal Reserve Bank of San Francisco.

Fernández-Villaverde, J., S. Hurtado, and G. Nuño (2023): "Financial Frictions and the Wealth Distribution," *Econometrica*, 91, 869–901, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA18180.

Friedman, M. (1953): "The Methodology of Positive Economics," in *Essays in Positive Economics*, ed. by M. Friedman, University of Chicago Press, 3–43.

Fuster, A., B. Hebert, and D. Laibson (2012): "Natural Expectations, Macroeconomic Dynamics, and Asset Pricing," *NBER Macroeconomics Annual*, 26, 1–48, publisher: The University of Chicago Press.

Gabaix, X. (2014): "A Sparsity-Based Model of Bounded Rationality *," *The Quarterly Journal of Economics*, 129, 1661–1710.

——— (2019): "Chapter 4 - Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, ed. by B. D. Bernheim, S. DellaVigna, and D. Laibson, North-Holland, vol. 2 of *Handbook of Behavioral Economics - Foundations and Applications 2*, 261–343.

Gertler, M., N. Kiyotaki, and A. Prestipino (2020): "A Macroeconomic Model with Financial Panics," *The Review of Economic Studies*, 87, 240–288.

Grandmont, J. M. (1977): "Temporary General Equilibrium Theory," *Econometrica*, 45, 535–572, publisher: [Wiley, Econometric Society].

Greenwood, R. and A. Shleifer (2014): "Expectations of Returns and Expected Returns," *The Review of Financial Studies*, 27, 714–746.

Gu, Z., M. Laurière, S. Merkel, and J. Payne (2024): "Global Solutions to Master Equations for Continuous Time Heterogeneous Agent Macroeconomic Models," ArXiv:2406.13726 [math].

GUVENEN, F. (2009): "An empirical investigation of labor income processes," *Review of Economic Dynamics*, 12, 58–79.

GUVENEN, F., S. OZKAN, AND J. SONG (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122, 621–660, publisher: The University of Chicago Press.

HAJDINI, I. (2023): "Mis-specified Forecasts and Myopia in an Estimated New Keynesian Model," *Working Paper*, institution: Federal Reserve Bank of Cleveland.

HANSEN, L. P. AND T. J. SARGENT (2008): *Robustness*, Princeton University Press, stu - student edition ed.

HOMMES, C. AND G. SORGER (1998): "CONSISTENT EXPECTATIONS EQUILIBRIA," *Macroeconomic Dynamics*, 2, 287–321.

JANSSENS, E. F. AND S. MCCRARY (2023): "Finite-State Markov-Chain Approximations: A Hidden Markov Approach," .

KAPLAN, G., K. MITMAN, AND G. L. VIOLANTE (2020): "The Housing Boom and Bust: Model Meets Evidence," *Journal of Political Economy*, 128, 3285–3345, publisher: The University of Chicago Press.

KHAN, A. AND J. K. THOMAS (2013): "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 121, 1055–1107, publisher: The University of Chicago Press.

KOHLHAS, A. N. AND A. WALTHER (2021): "Asymmetric Attention," *American Economic Review*, 111, 2879–2925.

KRUSELL, P., T. MUKOYAMA, AND A. ŞAHIN (2010): "Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations," *The Review of Economic Studies*, 77, 1477–1507.

KRUSELL, P. AND A. A. SMITH (2006): "Quantitative Macroeconomic Models with Heterogeneous Agents," in *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, ed. by R. Blundell, T. Persson, and W. K. Newey, Cambridge: Cambridge University Press, vol. 1 of *Econometric Society Monographs*, 298–340.

KRUSELL, P. AND J. SMITH, ANTHONY A. (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896, publisher: The University of Chicago Press.

KUCHLER, T. AND B. ZAFAR (2019): "Personal Experiences and Expectations about Aggregate Outcomes," *The Journal of Finance*, 74, 2491–2542, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12819.

LUCAS, R. E. (1972): "Expectations and the neutrality of money," *Journal of Economic Theory*, 4, 103–124.

——— (1976): "Econometric policy evaluation: A critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 19–46.

LUDVIGSON, S. C. AND A. MICHAELIDES (2001): "Does Buffer-Stock Saving Explain the

Smoothness and Excess Sensitivity of Consumption?" *American Economic Review*, 91, 631–647.

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

——— (2016): "Learning from Inflation Experiences *," *The Quarterly Journal of Economics*, 131, 53–87.

Mankiw, N. G. and R. Reis (2002): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve*," *The Quarterly Journal of Economics*, 117, 1295–1328.

Mankiw, N. G., R. Reis, and J. Wolfers (2003): "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual*, 18, 209–248, publisher: The University of Chicago Press.

Marcet, A. and T. J. Sargent (1989): "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," *Journal of Political Economy*, 97, 1306–1322, publisher: The University of Chicago Press.

Maćkowiak, B., F. Matějka, and M. Wiederholt (2023): "Rational Inattention: A Review," *Journal of Economic Literature*, 61, 226–273.

Midrigan, V. (2011): "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations," *Econometrica*, 79, 1139–1180, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA6735.

Molavi, P. (2024): "Simple Models and Biased Forecasts," ArXiv:2202.06921 [econ].

Molavi, P., A. Tahbaz-Salehi, and A. Vedolin (2024): "Model Complexity, Expectations, and Asset Prices," *The Review of Economic Studies*, 91, 2462–2507.

Moll, B. (2025): "The Trouble with Rational Expectations in Heterogeneous Agent Models: A Challenge for Macroeconomics," ArXiv:2508.20571 [econ].

Nagel, S. and Z. Xu (2022): "Asset Pricing with Fading Memory," *The Review of Financial Studies*, 35, 2190–2245.

Pischke, J.-S. (1995): "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica*, 63, 805–840, publisher: [Wiley, Econometric Society].

Reiter, M. (2010): "Approximate and Almost-Exact Aggregation in Dynamic Stochastic Heterogeneous-Agent Models," Number: 258 Place: Wien Publisher: Institut für Höhere Studien.

Rouwenhorst, K. G. (1995): "Asset Pricing Implications of Equilibrium Business Cycle Models," in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley, Princeton University Press, 294–330.

Sargent, T. J. (1996): "Expectations and the nonneutrality of Lucas," *Journal of Monetary Economics*, 37, 535–548.

——— (2008): "Evolution and Intelligent Design," *American Economic Review*, 98, 5–37.

Sims, C. A. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50, 665–690.

TAUCHEN, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20, 177–181.

YOUNG, E. R. (2010): "Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm and non-stochastic simulations," *Journal of Economic Dynamics and Control*, 34, 36–41.

Appendix

A The Krusell and Smith (1998) Parametric RPE

This appendix first defines the RPE that KS compute with their method, which approximates FIRE. Then, I define a parametric RPE where households forecast prices using prices. I define the equilibria for the environment outlined in Section 4.1, which only differs from the original KS environment in the evolution of the idiosyncratic shock: I assume that the distribution over ε' does not depend on $\{\theta, \theta'\}$, while KS do. This is without loss for the purposes of this appendix, which are to (1) illustrate that the equilibrium that KS compute is an RPE, (2) show that an RPE where agents forecast over prices is numerically very close, and (3) show that prices do not follow a Markov process.[43]

A.1 Recursive Problem and Equilibrium Definition

KS solve the equilibrium where agents observe capital (the first moment of the wealth distribution) and aggregate productivity, and forecast next period's aggregate capital with a log-linear perceived law of motion. They use the pricing functions $w(\cdot)$ and $r(\cdot)$ to form expectations over wages and the rental rate of capital respectively, given some future capital and productivity. Their recursive problem solves

$$
\begin{aligned}
v(a, \varepsilon; K, \theta) &= \max_{c, a'} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_{\pi,\Pi} \left[v\left(a', \varepsilon'; K', \theta'\right) \mid \varepsilon, \theta \right] \right\} \qquad (11)\\
\text{s.t.} \quad c + a' &= w(K, \theta)\, \varepsilon + \left[1 + r(K, \theta) - \delta\right] a, \\
a' &\geq \underline{a}, \\
\log K' &= \psi_0(\theta) + \psi_1(\theta) \log K.
\end{aligned}
$$

where the $\mathbb{E}_{\pi,\Pi}[\cdot \mid \varepsilon, \theta]$ operator highlights that expectations are formed over ε' and θ' using π and Π, so there is no uncertainty over aggregate equilibrium variables, as K' is assumed to follow a deterministic path conditional on $\{K, \theta\}$. That is, households only consider their point forecast when making expectations. Lastly, I use the discrete space of KS for θ, where $\theta \in \{\theta_l, \theta_h\}$.

While KS explain this recursive problem, they don't define the equilibrium that comes out of it. Their reason is that it approximates very well the FIRE equilibrium that they do define. However, it is instructive for our purposes to define the equilibrium that they compute. Let $p := \{w, r\}$ as in Section 4.2. For $K \in \boldsymbol{K}$, let $\mu \in \Delta\left(\boldsymbol{K}^2\right)$ denote the joint ergodic distribution of K and K', while $\mu_K \in \Delta(\boldsymbol{K})$ denotes the ergodic distribution of K.

Definition 4. *The Krusell Smith Restricted Perceptions Equilibrium (KSRPE) is a set of v, g, p, Ψ, μ and $\{\psi_i(\theta)\}_{i,\theta}$ such that:*

[43] (1) follows directly, and I have checked that (2) and (3) extend to the original KS process for ε, results available upon request.

1. *Given the pricing functions p and the perceived law of motion $\{\psi_i(\theta)\}_{i,\theta}$, the policy g is optimal and v is the resulting value function in the Bellman equation 11.*

2. *The pricing functions p satisfy firm optimality* (6)*.*

3. *Labor, capital and goods markets clear, satisfying* (7)*.*

4. *The law of motion Ψ is generated by $\{g, \pi\}$.*

5. *The distribution μ is induced by $\{\Psi, \Pi\}$.*

6. *For all θ,*

$$
\begin{aligned}
\psi_1(\theta) &= \frac{\mathbb{C}_\mu\left[\log K', \log K \mid \theta\right]}{\mathbb{V}_{\mu_K}\left[\log K \mid \theta\right]}, \\
\psi_0(\theta) &= \left[1 - \psi_1(\theta)\right] \mathbb{E}_{\mu_K}\left[\log K \mid \theta\right].
\end{aligned}
$$

In this definition, $\mathbb{C}_\mu[\cdot \mid \theta]$ refers to the conditional covariance under μ, given θ. In the KS method the economy is simulated for many periods, and $\log K'$ is regressed on $\log K$ in simulated data. The coefficients of the perceived law of motion in the KSRPE correspond to the limit of this OLS regression as the simulation length goes to infinity. In this sense, KS compute the RPE we have just defined.

The reader can note 2 differences with the NRPE from Section 4.2. First, in the KSRPE the perceived law of motion is over $\{K, \theta\}$, while in the NRPE it is over $\{w, r\}$. As we saw in that section, it is possible to show an equivalence between the NRPE and a generalization of the KSRPE. In the next subsection of this appendix, I show that the KSRPE is very close to a parametric RPE where the law of motion is over $\{w, r\}$.

Second, the KSRPE has a parametric perceived law of motion which produces a point forecast, instead of a distribution over possible outcomes. We learned from KS that these restrictions matter very little. I compute the KSRPE in the environment of this paper with the KS algorithm and the calibration of KS, and find an R^2 from the regression on simulated data to be ≈ 0.999998.

A.2 An RPE in Prices

This paper's choice of a nonparametric version of an RPE is motivated by many reasons, outlined in Section 3.3. However, it is still possible to define a parametric RPE over objects such as prices. In order to do so, I first revisit the point made in Section 4.3 that the prices are informationally equivalent to the KS aggregate states. This is made explicit by rearranging the firm optimality conditions:

$$
\begin{aligned}
K &= \frac{w}{r}\left(\frac{\alpha}{1-\alpha}\right) H, \\
\theta &= \left(\frac{w}{1-\alpha}\right)^{1-\alpha}\left(\frac{r}{\alpha}\right)^{\alpha},
\end{aligned}
\tag{12}
$$

which makes it clear that any household observing $\{w, r\}$ can back out $\{K, \theta\}$.[44] If they can infer the sequence of θ, then they can learn Π from a time series of $\{w, r\}$, and perfectly so in the asymptotic limit. I use this result when considering agents who solve the following problem:

$$
\begin{aligned}
v(a, \varepsilon; w, r) = \max_{c, a'} & \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_{\pi, \Pi} \left[v(a', \varepsilon'; w', r') \mid \varepsilon, w, r \right] \right\} \qquad (13) \\
\text{s.t.} \quad c + a' &= w\varepsilon + [1 + r - \delta] a, \\
a' &\geq \underline{a}, \\
\log w' &= \phi_0^w(\theta(w,r)) + \phi_1^w(\theta(w,r)) \log w + \phi_2^w(\theta'), \\
\log r' &= \phi_0^r(\theta(w,r)) + \phi_1^r(\theta(w,r)) \log r + \phi_2^r(\theta'), \\
\mathbb{P}(\theta' = x \mid \theta = \theta(p)) &= \Pi(x \mid \theta(p)).
\end{aligned}
$$

Relative to (11), there are now (naturally) two perceived laws of motion: one for wages and one for rental rate of capital. The coefficients for these regressions depend on the object $\theta(w, r)$ that is constructed using (12). Effectively they are backing out θ from the observed prices. In addition they must necessarily be stochastic: agents must expect their prices to change depending on the future's aggregate productivity shock. Since they know Π and they can back out θ, this can be dealt with easily by including the terms $\phi_2^w(\theta')$ and $\phi_2^r(\theta')$.

I will define the equilibrium that results from this recursive problem in a style more similar to Definition 3, since it is more compact. For $p \in P$, $\mu \in \Delta(P^2)$ denotes the joint ergodic distribution of p and p'. Let $Q_\phi : P \to \Delta(P)$ denote the perceived law of motion in (13) with parameters $\phi = \left\{\phi_j^i(\theta)\right\}_{ij\theta}$ for $i = \{w, r\}$ and $j = \{0, 1, 2\}$.

Definition 5. *The Restricted Perceptions Equilibrium in Prices (RPEP) is a set* $\{v, g, p, \Psi, \mu, Q_\phi\}$ *such that:*

1. *Given Q_ϕ, the policy g is optimal and v is the resulting value function in the Bellman equation 13.*

2. *Prices p satisfy firm optimality (6).*

3. *Conditions 3, 4 and 5 of the KSRPE hold.*

4. *Q_ϕ minimizes the least square errors under μ.*

I solve for this equilibrium with the KS algorithm, while running the regressions on prices when updating the perceived law of motion. I get an $R^2 \approx 0.999999$, marginally higher than in the KSRPE. This is a numerical confirmation of the intuition outlined above: prices are informationally equivalent to capital and productivity. In the RPEP, I run the aggregate capital regressions that form the perceived law of motion of the KSRPE, and I find practically identical dynamics, as shown in Table 2.[45]

[44] Note that H is constant in this economy, which makes the exposition simpler. In KS it is a function of θ, but the point still stands if we write $H(\theta)$ in the first equation above.

[45] Theoretically, there is no guarantee to find the *exact* same dynamics. This is because the minimization

Table 2: Comparison of Dynamics

	$\psi_0(\theta_l)$	$\psi_1(\theta_l)$	$\psi_0(\theta_h)$	$\psi_1(\theta_h)$
KSRPE	0.1225	0.9655	0.1273	0.9646
RPEP	0.1224	0.9655	0.1272	0.9647

Lastly, in the RPEP I can check whether looking at longer lags of prices provides more information to forecast next period's prices. That is, we can check whether prices are a Markov process. The intuition discussed by Moll (2025) is confirmed numerically: adding one more lag to the perceived law of motion in (13) increases the fit of the price regression to $R^2 \approx 0.9999999999$. This means that in this equilibrium, short beliefs bite too, at least in theory. In this economy, the difference in fit is negligible only because the fit with 1 lag of prices is so high. In other economies, this might not be the case.

B Nonparametric Generalization of Krusell and Smith (1998)

I start by stating the recursive problem of the agent. It will be useful to denote $x = \{K, \theta\}$ and frame it as:

$$v_x(a, \varepsilon; x) = \max_{c, a'} \left\{ \frac{c^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_{\pi, Q_x} \left[v_x(a', \varepsilon'; x') \mid \varepsilon; x \right] \right\} \tag{14}$$
$$\text{s.t.} \quad c + a' = w(x)\varepsilon + [1 + r(x) - \delta] a,$$
$$a' \geq \underline{a},$$
$$x' \sim Q_x(x)$$

We can see that relative to (11), the only change is on the expectations over K', which is now a random variable with a joint distribution with θ'. For $x \in X$, let $\mu_x \in \Delta(X^2)$ denote the joint ergodic distribution of $\{K, K', \theta, \theta'\}$. This allows us to define the necessary equilibrium:

Definition 6. *The Krusell Smith Nonparametric Restricted Perceptions Equilibrium (KSNRPE) is a set $\{v, g, p, \Psi, \mu, Q\}$ such that:*

1. *Given the pricing functions p and the perceived law of motion Q, the policy g is optimal and v is the resulting value function in the Bellman equation* (14).
2. *The pricing functions p satisfy firm optimality* (6).
3. *Labor, capital and goods markets clear, satisfying* (7).
4. *The law of motion Ψ is generated by $\{g, \pi\}$.*

of least square errors is defined over different variables in both equilibria, which might lead to small differences. There might also be small differences due to computational limits to accuracy. I do not explore this issue further since I provide a theoretical proof of equivalence for the NRPE in Appendix B.

5. *The distribution μ is induced by $\{\Psi, \Pi\}$.*
6. *For any measurable $\mathcal{X}, \mathcal{X}'$ in the Borel sigma-algebra of X:*

$$\int_{\mathcal{X}} Q_x\left(\mathcal{X}' \mid x\right) \mu_x\left(dx, X\right) = \mu_x\left(\mathcal{X} \times \mathcal{X}'\right)$$

We are now in a position to prove the equivalence result.

B.1 Proof of Proposition 1

I will start by showing that for any NRPE, there exists a KSNRPE with the same Ψ. The proof will proceed by guess and verify. Guess that the law of motion for both economies is indeed Ψ. This means that, given the same sequence of shocks, the aggregate variables in the economy must be identical. Therefore μ and μ^{KS} must be mutually consistent in the sense that, for any $\mathcal{X}$ and $\mathcal{X}'$, we have:

$$\mu_x\left(\mathcal{X} \times \mathcal{X}'\right) = \mu\left(f\left(\mathcal{X}\right) \times f\left(\mathcal{X}'\right)\right),$$

where $f\left(\cdot\right)$ denotes the injective mapping from $\{K, \theta\}$ to $\{w, r\}$ provided by the firm optimality conditions. Using this relation, the belief consistency conditions imply a similar equivalence between Q and Q_x:

$$Q_x\left(\mathcal{X} \mid x\right) = Q\left(f\left(\mathcal{X}\right) \mid f\left(x\right)\right). \tag{15}$$

Lastly, recall the recursive problem in the NRPE, while using $p = \{w, r\}$:

$$\begin{aligned} v\left(a, \varepsilon; p\right) = \max_{c, a'} & \left\{\frac{c^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_{\pi, Q}\left[v\left(a', \varepsilon'; p'\right) \mid \varepsilon; p\right]\right\} \\ \text{s.t.} \quad & p' \sim Q\left(p\right). \end{aligned}$$

Do the change of variables $p = f\left(x\right)$ and $p' = f\left(x'\right)$:

$$\begin{aligned} v\left(a, \varepsilon; f\left(x\right)\right) = \max_{c, a'} & \left\{\frac{c^{1-\gamma}}{1-\gamma} + \beta \mathbb{E}_{\pi, Q}\left[v\left(a', \varepsilon'; f\left(x'\right)\right) \mid \varepsilon; f\left(x\right)\right]\right\} \\ \text{s.t.} \quad & f\left(x'\right) \sim Q\left(f\left(x\right)\right). \end{aligned}$$

Because of (15), and adding the budget constraint, this problem can be stated as:

$$
\begin{aligned}
v\left(a,\varepsilon;f\left(x\right)\right) = \max_{c,a'} & \left\{\frac{c^{1-\gamma}}{1-\gamma} + \beta\mathbb{E}_{\pi,Q}\left[v\left(a',\varepsilon';f\left(x'\right)\right) \mid \varepsilon;f\left(x\right)\right]\right\} \\
\text{s.t.} \quad & c + a' = w\left(x\right)\varepsilon + \left[1 + r\left(x\right) - \delta\right]a, \\
& a' \geq \underline{a}, \\
& x' \sim Q_x\left(x\right)
\end{aligned}
$$

One glance at (14) allows us to conclude that:

$$
v_x\left(a,\varepsilon;x\right) = v\left(a,\varepsilon;f\left(x\right)\right)
$$

Therefore, when faced by the same aggregate states x, agents in these equilibria face the same problem. The implication is that the policy functions satisfy $g_x\left(a,\varepsilon;x\right) = g\left(a,\varepsilon;f\left(x\right)\right)$. And if the policy functions are the same, then it must be that Ψ is the same in these two economies, verifying the guess. Since $f\left(\cdot\right)$ is injective, a completely analogous argument can be made to reach the conclusion that $g_x\left(a,\varepsilon;f^{-1}\left(p\right)\right) = g\left(a,\varepsilon;p\right)$ as well, thus showing that for any KSNRPE, there exists a NRPE with the same Ψ. □

C Markov Chain Approximation

This section describes a novel computational method to approximate a Markov chain from a multivariate time series $\{s_t\}_{t=1}^{T}$. The intention is to use this method on the simulated data from the algorithm outlined in Section 3.3.1. The method follows the logic of a histogram approach, which would compute the shares of transitions from one region of the state space (the continuous space where the time series lives) to another region, where each region would have one element to conform the discrete space of the Markov chain. Then, the transition shares would correspond to the transition probabilities. Denote the discrete space of the Markov chain as the grid $\boldsymbol{s}$.

However, it is a generalization from such an approach, in that I consider observations s_t closer to the a given gridpoint to have a higher weight when computing transition probabilities, where the weight is computed with kernel densities. This achieves three objectives: first it increases the accuracy of the approximation, since the histogram regions might be wide and thus observations far from the grid might have undue influence. Second, it allows for better stability in the algorithm from Section 3.3.1.[46] Third, it scales better as the number of gridpoints increases and the region becomes narrower, in which case the histogram method would require an increasingly large sample T to cover each region accurately. In any case, the histogram approach is nested by choosing a uniform kernel with a given bandwidth, as will be clear below.

In addition, I choose the gridpoints to match the unconditional first and second moments exactly, by computing the transition probabilities for data that has been whitened,

[46]The reason being that between iterations in the algorithm, small movements in $\{s_t\}_{t=1}^{T}$ might result in large changes in the transition probabilities if points are close to the borders of the regions.

in the sense of having zero means and a diagonal covariance matrix. In this whitened space, the researcher must choose a grid $\tilde{\mathbf{s}}$. As an example, I use the tensor grid generated by the Rouwenhorst (1995) method, which in the whitened space only needs the number of gridpoints.[47]

The approximation method for a given bandwidth matrix h (the choice of which I discuss below) proceeds as follows:

1. Whiten the data to obtain $\{\tilde{s}_t\}_{t=1}^{T}$. That is, if μ_s is the sample mean and Σ_s the sample covariance matrix of $\{s_t\}_{t=1}^{T}$, for every t:

$$\tilde{s}_t = (s_t - \mu_s)\,\Sigma_s^{-\frac{1}{2}}$$

2. For each gridpoint i and time t, compute the weight ω_{it} of simulated data by evaluating the distance in a kernel with bandwidth matrix h:

$$\omega_{it} = K_h\left(\tilde{s}_t - \tilde{\mathbf{s}}^i\right)$$

   I discuss the form and choice of kernel below.

3. Compute transition measures from gridpoint i to j as:

$$\hat{Q}_{ij} = \frac{\sum_{t=1}^{T-1} \omega_{it}\omega_{jt+1}}{\sum_{t=1}^{T-1} \omega_{it}}$$

4. Normalize the transition measures to compute transition probabilities:

$$Q_{ij} = \frac{\hat{Q}_{ij}}{\sum_j \hat{Q}_{ij}}$$

5. Given probabilities Q, compute the grids $\mathbf{s}$ that exactly match μ_s and Σ_s. Note that this produces a non-tensor grid if Σ_s is not diagonal.

An example of a kernel that works well in my application is the Gaussian one, given by $K_h(u) = \exp\left(-u'h^{-1}u\right)$. The continuous support over the real line means that, if it so happens during the algorithm that no points land near a given gridpoint, a transition probability can still be computed.[48] This is a practical advantage over, say, an Epanechnikov kernel. At the same time, the thin tails of the Gaussian distribution ensure that observations significantly far away from the gridpoint have very little weight in comparison to observations that are closer. One may also notice that the histogram approach is nested by the uniform kernel $K_h(u) = \prod_{d=1}^{D} \mathbb{I}\left(|u_d| \leq h_d\right)(2h_d)^{-1}$, where

[47] For N gridpoints, this grid is even-spaced from $-\sqrt{N-1}$ to $\sqrt{N-1}$.

[48] It will not matter for equilibrium dynamics if the transition probabilities starting from this gridpoint dynamics are inaccurate, because that region of the state space is not visited anyway.

Table 3: Moments in Levels

	$\sigma(Y)$	$\sigma(C)$	$\sigma(I)$	$cor(C,Y)$	$cor(I,Y)$	$\rho(Y)$	$\rho(C)$	$\rho(I)$
FIRE	1.04	0.58	1.62	0.97	0.99	0.73	0.73	0.72
NB	1.04	0.41	2.01	0.6	0.98	0.72	0.69	0.72
Data	1.63	1.37	7.10	0.79	0.82	0.79	0.72	0.78

Table 4: Moments in Differences

	$\sigma(\Delta Y)$	$\sigma(\Delta C)$	$\sigma(\Delta I)$	$cor(\Delta C,\Delta Y)$	$cor(\Delta I,\Delta Y)$	$\rho(\Delta Y)$	$\rho(\Delta C)$	$\rho(\Delta I)$
FIRE	0.77	0.43	1.21	0.97	0.99	-0.07	-0.07	-0.07
NB	0.77	0.32	1.5	0.56	0.97	-0.07	-0.1	-0.07
Data	1.06	1.04	4.71	0.73	0.76	0.06	-0.15	0.13

$\mathbb{I}(\cdot)$ is the indicator function and d denotes the different dimensions evaluated, so u_d are the elements of the vector u, and analogously for h_d.

The choice of h is highly relevant to the accuracy of the approximation, and can be understood in terms of a bias-variance tradeoff. If the bandwidth is too small, then only observations that are very close to the gridpoints matter. This means that we require longer simulations for accuracy, which is computationally costly. If the bandwidth is too large, then observations far away from the gridpoints matter too much, which gives us distorted estimates of the transition probabilities.

I propose the following method: consider diagonal h and choose the elements to match the persistence of each component of the multivariate time series $\{s_t\}_{t=1}^{T}$ by running a simulated method of moments. The idea is that a higher h tends to decrease persistence, because all observations are weighted relatively equally, and so the method produces zero persistence in the limit as $h \to \infty$. In practice, the numerical match is often very good. For example, when solving the equilibrium of Section 4.2, the simulated equilibrium persistence of w_t is given by 0.9955, while the Markov chain approximation produces a persistence of 0.9956. For r_t, these numbers are 0.9761 and 0.9762 respectively. This match in persistence, alongside the perfect match in unconditional moments, suggest that the dynamics of $\{s_t\}_{t=1}^{T}$ can be well approximated by this method.

D Additional Business Cycle Moments

This section shows additional business cycle moments from the equilibria computed in Sections 4.2 and 4.3. The moments are calculated following the same procedures as outlined in Section 4.5. Table 3 outlines the moments in levels, while Table 4 shows the same moments for first differences. It can be appreciated that looking at levels or differences produces the same qualitative takeaways.

The most apparent fact is that both equilibria fall short of the output volatility that is observed in the data at a business cycle frequency. This is due to the high persistence of the θ shock, which generates volatility at a frequency lower than that of usual business cycles,[49] and the absence of other shocks of lower persistence. Consequently the volatility of both consumption and investment is also too low relative to the data, but the relative magnitudes are better matched by narrow beliefs, as shown explicitly in Section 4.5.

In addition, we can observe that consumption is overly procyclical in FIRE, and not procyclical enough with narrow beliefs. Investment is consistently overly procyclical in both equilibria. The autocorrelations of each series in levels for both equilibria are roughly in line with the data for levels, but they are somewhat off for differences, again suggesting the absence of some aggregate shocks with lower persistence.

[49] If I calculate the data moments subtracting a log-linear trend, both equilibria actually overshoot the volatility of output relative to the data, and more so for the narrow beliefs equilibrium.