Uncertainty Traps*

Pablo Fajgelbaum — UCLA and NBER
Edouard Schaal — NYU
Mathieu Taschereau-Dumouchel — Wharton

May 10, 2016

Abstract

We develop a theory of endogenous uncertainty and business cycles in which short-lived shocks can generate long-lasting recessions. In the model, higher uncertainty about fundamentals discourages investment. Since agents learn from the actions of others, information flows slowly in times of low activity and uncertainty remains high, further discouraging investment. The economy displays uncertainty traps: self-reinforcing episodes of high uncertainty and low activity. While the economy recovers quickly after small shocks, large temporary shocks may have long-lasting effects on the level of activity. The economy is subject to an information externality but uncertainty traps may remain in the efficient allocation. Embedding the mechanism in a standard business cycle framework, we find that endogenous uncertainty increases the persistence of large recessions and improves the performance of the model in accounting for the Great Recession.

JEL Classification: E32, D80

*We thank the editor Robert Barro and two anonymous referees for valuable suggestions. Liyan Shi and Chunzan Wu provided superb research assistance. We are grateful to Andrew Abel, Andrew Atkeson, Michelle Alexopoulos, Jess Benhabib, Harold Cole, Michael Evers, João Gomes, William Hawkins, Patrick Kehoe, Lars-Alexander Kuehn, Ali Shourideh, Laura Veldkamp, Pierre-Olivier Weill and seminar participants for useful comments. Corresponding author: Edouard Schaal, 19 W. 4th Street, 6FL, New York, NY 10012, *edouard.schaal@nyu.edu*.

1 Introduction

We develop a theory of endogenous uncertainty and business cycles. The theory combines two forces: higher uncertainty about economic fundamentals deters investment, and uncertainty evolves endogenously because agents learn from the actions of others. The unique rational expectation equilibrium of the economy features *uncertainty traps*: self reinforcing episodes of high uncertainty and low economic activity that cause recessions to persist. Because of uncertainty traps, large but short-lived shocks can generate long-lasting recessions. We first build and characterize a model that only includes the essential features that give rise to uncertainty traps. Then, we embed these features into a standard real business cycle model and quantify the impact of endogenous uncertainty during the Great Recession.

In the model, firms decide whether to undertake an irreversible investment whose return depends on an imperfectly observed fundamental that evolves randomly according to a persistent process. Firms are heterogeneous in the cost of undertaking this investment and hold common beliefs about the fundamental. Beliefs are regularly updated with new information, and, in particular, firms learn by observing the return on the investment of other producers. We define uncertainty as the variance of these beliefs.

This environment naturally produces an interaction between beliefs and economic activity. Firms are more likely to invest if their beliefs about the fundamental have higher mean, but also if they have smaller variance (lower uncertainty). At the same time, the laws of motion for the mean and variance of beliefs depend on the investment rate. In particular, when few firms invest, little information is released, so uncertainty rises.

The key feature of the model is that this interaction between information and investment leads to uncertainty traps, formally defined as the coexistence of multiple stationary points in the dynamics of uncertainty and economic activity. Without shocks, the economy converges to either a high regime (with high economic activity and low uncertainty) if the current level of uncertainty is sufficiently low, or to a low regime (with low activity and high uncertainty) if the current level of uncertainty is sufficiently high. Because of the presence of these multiple stationary points, the economy exhibits non-linearities in its response to shocks: starting from the high regime, it quickly recovers after small temporary shocks, but it may shift to the low-activity regime after a large temporary shock. Once it has fallen in the low regime, only a large enough positive shock can push the economy back to the high-activity regime.

An important feature of the model is that, despite the presence of uncertainty traps, there is a unique recursive competitive equilibrium. That is, multiplicity of stationary points does not mean multiplicity of equilibria. Therefore, unlike other macro models with complementarities, there is no room in our model for multiple equilibria or sunspots.[1]

The model features an inefficiently low level of investment because agents do not internalize the effect of their actions on public information. This inefficiency naturally creates room for welfare-

[1]For recent examples of business cycle models with multiple equilibria see Farmer (2013), Kaplan and Menzio (2013), Benhabib et al. (2015) and Schaal and Taschereau-Dumouchel (2015).

enhancing policy interventions. We therefore study the problem of a constrained planner that is subject to the same informational constraints as private agents. The socially constrained-efficient allocation can be implemented with a subsidy to investment. But, perhaps surprisingly, the optimal policy does not necessarily eliminate uncertainty traps. Therefore, while policy interventions are desirable, they do not eliminate the adverse feedback loop between uncertainty and economic activity.

To evaluate the quantitative implications of uncertainty traps, we embed the key features of the baseline model into a standard general equilibrium framework and then compare its predictions with an RBC model and the data. To isolate the impact of endogenous movements in uncertainty, we also compare our full model to a "fixed θ-uncertainty" version in which uncertainty about the fundamental productivity θ is fixed over time. We discipline the key parameters of the model, those that determine option-value effects and the evolution of uncertainty, by targeting moments from the distribution of uncertainty about real GDP growth from the Survey of Professional Forecasters (SPF).

We first show that our calibrated model performs as well as the RBC and fixed θ-uncertainty models in terms of traditional business cycle moments. Therefore, incorporating endogenous uncertainty in a standard business cycle model does not impair its ability to predict well-known patterns of business cycle data.

Then, we demonstrate that the non-linearities generated by uncertainty traps, studied in the baseline theory, are active in the calibrated model. Specifically, we compute the economy's response to one-period negative shocks to beliefs of different magnitudes. We find that i) recessions are longer and deeper under the full model than under the fixed θ-uncertainty model, and ii) the difference between both models is more important for large shocks than for small ones. In response to a -1% shock, the ensuing recession is 22% deeper (in terms of the peak-to-trough fall in output) and 40% longer (in terms of quarters until the economy has recovered half of the peak-to-trough fall in output) in the full model than in the fixed θ-uncertainty model. However, in response to a larger -5% shock, the recession is 35% deeper and 66% longer in the full model than in the fixed θ-uncertainty model. Therefore, in the calibrated model, the endogenous uncertainty mechanism, whose impact is captured by the difference between the full and fixed θ-uncertainty models, makes recessions deeper and longer for shocks of any magnitude, but relatively more so for larger shocks.

Finally, our main quantitative exercise evaluates the predictions of our calibrated model for past U.S. recessions. Since our mechanism provides amplification and persistence to large shocks, we expect that it might help explain particularly severe recessions observed in the data. We therefore investigate the largest recession in our sample, the Great Recession. To do so, we feed each of the three models (our full model, the RBC model, and the fixed θ-uncertainty model) with the observed TFP series and signals such that each model replicates the time series of forecasts about output growth from the SPF during the first part of the recession. We then contrast each model's response with the data.

Our main quantitative finding is that, during the Great Recession, our model generates declines in output, consumption, employment, and investment which are clearly more protracted, and closer to patterns observed in the data, than what the alternative models predict. Endogenous uncertainty adds 1.8 percentage points in terms of recession's depth and slows the recovery by about two years relative to the fixed θ-uncertainty model.[2] The corresponding numbers are 5.2 percentage points and five years when comparing to the RBC model. We also evaluate the performance of the three models against the data in terms of one key statistic that summarizes both the depth and the length of the recession: the cumulative output loss between the start of the recession and 2015. We find that our model generates 93% of the Great Recession's cumulative output lost, relative to the 70% and 30% generated by the fixed θ-uncertainty and RBC models, respectively. Reassuringly, the model also generates patterns for the evolution of uncertainty about output growth that are roughly consistent with the data.

We demonstrate the robustness of these conclusions by replicating this exercise under alternative assumptions about the shocks hitting the economy, the source of the TFP data, the detrending strategy, and the preferences of the household. In each case, we find that the model with endogenous uncertainty performs better than its alternatives. To make sure that the full model does not generate counterfactual amounts of persistence for milder recessions, we also replicate the second largest recession in our sample, the 1981-1982 recession, which was characterized by a relatively rapid recovery. We find that our model behaves similarly to the RBC and to the fixed θ-uncertainty models in that case. We conclude that the inclusion of uncertainty traps in a standard macroeconomic model of business cycles improves its performance during the Great Recession and that it leads to similar predictions than standard models for smaller recessions.

The remainder of the introduction contains the literature review followed by a discussion of our notion of uncertainty and its business cycle properties. The paper is then structured as follows. Section 2 presents the baseline model and the definition of the recursive equilibrium. Section 3 characterizes the investment decision of an individual firm and demonstrates the existence and uniqueness of the equilibrium. Section 4 shows the existence of uncertainty traps, examines the non-linearities that they generate, and characterizes the planner's problem. Section 5 describes the quantitative model, shows how uncertainty traps influence the response of the economy to shocks and compares the dynamic properties of our model to an RBC model, a fixed θ-uncertainty model and the data over the Great Recession. Section 6 concludes. The full statement of the proposition and the proofs can be found in the appendix.

1.1 Relation to the Literature

The theory is motivated by an empirical literature that investigates the impact of uncertainty on economic activity using VARs, as in Bloom (2009) and Bachmann et al. (2013), or using instrumental variables, as in Carlsson (2007), and finds that increases in uncertainty typically slow

[2]This measure refers to the time the economy takes to recover 20% of its peak-to-trough decline. We use the 20% threshold instead of the usual half-life since, in the data, detrended output has only recovered about 20% of its peak-to-trough decline by the end of our sample in 2015.

down economic activity. It also relates to the uncertainty-driven business cycle literature that analyzes the impact of uncertainty through real option effects as in Bloom (2009), Bloom et al. (2012), Bachmann and Bayer (2013) and Schaal (2015), or through financial frictions as in Arellano et al. (2012) and Gilchrist et al. (2014).[3]

Our analysis also relates to a theoretical literature in macroeconomics that studies environments with learning from market outcomes such as Veldkamp (2005), Ordoñez (2009) and Amador and Weill (2010). Closely related to our paper is the analysis of Van Nieuwerburgh and Veldkamp (2006). They focus on explaining business-cycle asymmetries in an RBC model with incomplete information in which agents receive signals with procyclical precision about the economy's fundamental. During recessions, agents discount new information more heavily and the mean of their beliefs recovers slowly. Their paper provides a theory of endogenous pessimism that can explain business cycle asymmetries. Our model introduces a similar learning environment in a model of irreversible investment under uncertainty in the spirit of Dixit and Pindyck (1994) and Stokey (2008). The resulting feedback loop between endogenous uncertainty and real option effects, specific to our approach, offers a novel propagation mechanism that can lead to persistent episodes of high uncertainty and low economic activity.

The interaction of endogenous uncertainty and real option effects in our model is also reminiscent of the literature on learning and strategic delays as in Lang and Nakamura (1990), Rob (1991), Caplin and Leahy (1993), Chamley and Gale (1994), Zeira (1994) and Chamley (2004). Our paper differs from this literature in its attempt to evaluate and quantify the role of uncertainty and delays in a standard business cycle framework. In a recent paper in which learning and economic activity interact, Straub and Ulbricht (2015) propose a theory of endogenous uncertainty in which financial constraints impede learning about firm-level fundamentals. Financial crises cause uncertainty to rise, leading to a further tightening of financial constraints that amplifies and propagates recessions.[4] In another recent paper considering the role of learning during the Great Recession, Kozlowski et al. (2015) suggest that the Great Recession was the result of an unlikely shock that caused agents to substantially revise their beliefs about the probability of lower-tail events. They find that the resulting increase in pessimism may account for part of the long-lasting downturn.

This paper is also related to the literature on fads and herding in the tradition of Banerjee (1992) and Bikhchandani et al. (1992) . Articles in that tradition consider economies with an unknown fixed fundamental and study a one-shot evolution towards a stable state, whereas we study the full cyclical dynamics of an economy that fluctuates between regimes.

The dynamics generated by the model, with endogenous fluctuations between regimes, is reminiscent of the literature on static coordination games such as Morris and Shin (1998, 1999) and the dynamic coordination games literature as Angeletos et al. (2007) and Chamley (1999). These papers study games in which a complementarity in payoffs leads to multiple equilibria under complete

[3]Another literature studying time-varying risk is the literature on rare disasters (Barro, 2006) and time-varying disaster risk as in Gabaix (2012) and Gourio (2012), and surveyed in Barro and Ursúa (2012).

[4]Some recent papers discuss alternative channels that give rise to endogenous volatility over the business cycle. See Bachmann and Moscarini (2011) and Decker and D'Erasmo (2016).

information. The introduction of strategic uncertainty through noisy observation of the fundamental leads to a departure from common knowledge that eliminates the multiplicity. In contrast, the complete-information version of our model does not feature multiplicity, and complementarity only arises under incomplete information through social learning. Uniqueness does not obtain through strategic uncertainty, but by limiting the strength of the complementarities.

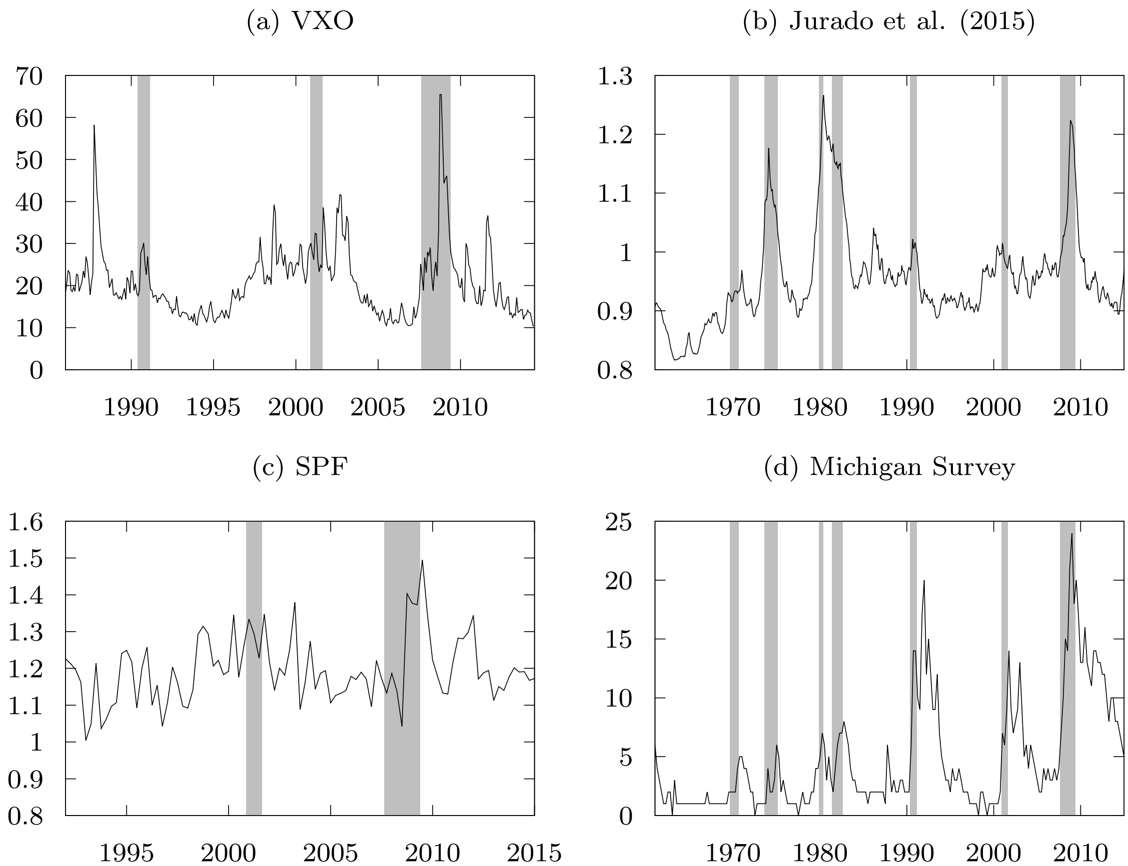
1.2 Bayesian Uncertainty and the Business Cycle

Throughout the paper, we adopt the concept of *Bayesian uncertainty*: in our theory, all agents have access to the same information $\mathcal{I}_t$ at time t and use Bayes' rule to form beliefs about the *fundamental* of the economy θ_t, which is, in our context, the aggregate productivity process. We define *uncertainty* as the variance $\text{Var}(\theta_t \mid \mathcal{I}_t)$ of the probability distribution that describes these common beliefs. In contrast, the uncertainty-driven business cycle literature that we referenced above defines uncertainty as time-varying volatility in exogenous aggregate or idiosyncratic variables.

These two definitions of uncertainty are related, but they are not identical. They are related because time-varying volatility may generate uncertainty about the future fundamentals of the economy, giving rise to Bayesian uncertainty. However, they are different because Bayesian uncertainty may fluctuate without the presence of time-varying volatility. In our model, the variance of beliefs varies over time through learning, while the volatility of exogenous variables is constant.

A basic and well known feature of the data which motivates our theory is that uncertainty increases during recessions. Instead of direct measures of time-varying volatility, we present, in Figure 1 , the evolution of four measures that capture our notion of Bayesian uncertainty to the extent that they reflect uncertainty in subjective beliefs.[5] Panel (a) shows the VXO, a measure of stock market volatility as perceived by market participants; Panel (b) shows the uncertainty measure proposed by Jurado et al. (2015), which captures a Bayesian notion of ex-ante forecast error in a statistical model of the macroeconomy; Panel (c) shows the standard deviation of the average perceived distribution of output growth from the Survey of Professional Forecasters (SPF); and Panel (d) shows the fraction of respondents who answer "uncertain future" as a reason for why it is a bad time to buy major household goods from the Michigan Survey of Consumers. While these series attempt to measure distinct objects, they all capture the notion of subjective uncertainty. All these measures support the key implication of our mechanism, that uncertainty rises during recessions. In the quantitative section of the paper, we use the SPF measure to calibrate and evaluate the performance of our model because it has a natural counterpart in our framework.

[5]The uncertainty-driven business cycle literature measures aggregate uncertainty by the conditional heteroskedasticity of various aggregates such as TFP (Bloom et al., 2012). Time-varying volatility in idiosyncratic variables is typically proxied by cross-sectional dispersions in sales growth rates (Bloom, 2009), output and productivity (Kehrig, 2011), prices (Vavra, 2014), employment growth (Bachmann and Bayer, 2014), or business forecasts (Bachmann et al., 2013). All these measures have been shown to be countercyclical. Since all agents have the same beliefs about θ, these cross-sectional measures are uninformative about uncertainty in our model.



Notes: (a) The CBOE's VXO series is a measure of market expectations of stock market volatility over the next 30 days constructed from S&P100 option prices. We present monthly averages of the series over 1986-2014. (b) Jurado et al. (2015) estimate a large-scale structural model with time-varying volatility on the US economy and use it compute an implied measure of ex-ante forecast error. The series we present corresponds to the H12 measure, i.e., an equal-weighted average of the 12-month ahead standard deviations over 132 macroeconomic series. (c) The SPF series is the standard deviation of the "mean probability forecast": an average of the probability distribution provided over forecasters, of one-year ahead output growth in percentage terms. (d) The Michigan Survey series correspond to the percent fraction of all respondents that reply "uncertain future" to the question why people are not buying large household items. Shaded areas correspond to NBER recessions.

Figure 1: Various measures of subjective uncertainty

2 Baseline Model

We begin by presenting a stylized model that only features the necessary ingredients to generate uncertainty traps. The intuitions from this simple model as well as the laws of motion governing the dynamics of uncertainty carry through to the extended model that we use for numerical analysis.

2.1 Population and Technology

Time is discrete. There is a fixed number of firms $\overline{N}$, chosen large enough that firms behave atomistically. Each firm $j \in \left\{1, \ldots, \overline{N}\right\}$ holds a single investment opportunity that produces output θ, common to all firms. We refer to θ as the economy's fundamental. and assume that it follows the autoregressive process

$$\theta' = \rho_\theta \theta + \varepsilon^\theta, \ \varepsilon^\theta \sim \text{iid } \mathcal{N}\left(0, \left(1 - \rho_\theta^2\right) \sigma_\theta^2\right), \tag{1}$$

where $0 < \rho_\theta < 1$ is the persistence of the process and σ_θ^2 the variance of its ergodic distribution. To produce, a firm must pay a fixed cost f, drawn each period from the continuous cumulative distribution F with mean μ_f and standard deviation σ_f. Once production has taken place, the firm exits the economy and is immediately replaced by a new firm holding an investment opportunity. This assumption ensures that the mass of firms in the economy remains constant.[6]

Upon investment, the firm receives the payoff θ. Firms have constant absolute risk-aversion,[7]

$$u(\theta) = \frac{1}{a}\left(1 - e^{-a\theta}\right),$$

where $a > 0$ is the coefficient of absolute risk aversion.

2.2 Timing and Information

Firms do not know the true value of the fundamental θ and decide whether to invest or not based on their beliefs. As time unfolds, they learn about θ in various ways. First, they learn from a public signal Z with precision $\gamma_z > 0$ observed at the end of each period,

$$Z = \theta + \varepsilon^Z, \ \varepsilon^z \sim \text{iid } \mathcal{N}(0, \gamma_z^{-1}). \tag{2}$$

This signal captures the information released by statistical agencies or the media. Second, agents acquire information through social learning. When firm j invests, a noisy signal about its return, $x_j = \theta + \varepsilon_j^x$, is sent to all firms.[8] The noise ε_j^x is normally distributed with precision $\gamma_x/\overline{N} > 0$,

[6] This assumption is made for tractability and is relaxed in the quantitative section.

[7] Here, agents can be thought of as entrepreneurs with risk averse preferences. In our quantitative model, firms use the representative household's stochastic discount factor.

[8] Social learning captures the idea that firms learn from each other about various common components that affect their revenues such as productivity, demand, regulations, etc. Social learning has been found to influence economic decisions in various contexts. Foster and Rosenzweig (1995) estimate a model of the adoption of high-yielding seeds in India and find it consistent with social learning. Guiso and Schivardi (2007) find that peer-learning effects matter

independent over time and across investors, but common to all observers.[9] We denote by $N \in \{0, \ldots, \overline{N}\}$ the endogenous number of firms that invest and $n = N/\overline{N}$ the fraction of investing firms. Because of the normality assumption, a sufficient statistic for the information provided by investing firms is the public signal

$$X \equiv \frac{1}{N} \sum_{j \in I} x_j = \theta + \varepsilon_N^X, \tag{3}$$

where I is the set of such firms, and

$$\varepsilon_N^X \equiv \frac{1}{N} \sum_{j \in I} \varepsilon_j^x \sim \mathcal{N}\left(0, (n\gamma_x)^{-1}\right).$$

Importantly, the precision $n\gamma_x$ of this signal increases with the fraction of investing firms n.

The timing of events is summarized in Figure 2.

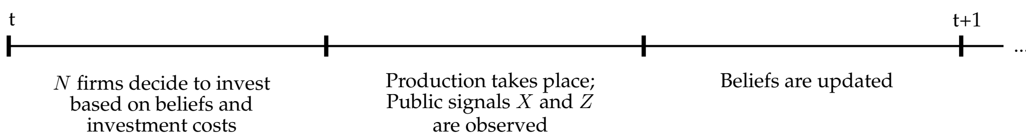


Figure 2: Timing of events

2.3 Beliefs

Under the assumption of a common initial prior, and because all information is public, beliefs are common across firms. In particular, there is no cross-sectional dispersion in beliefs. The normality assumptions about the signals and the fundamental imply that beliefs are also normally distributed

$$\theta \mid \mathcal{I} \sim \mathcal{N}\left(\mu, \gamma^{-1}\right),$$

where $\mathcal{I}$ is the information set at the beginning of the period. The mean of the distribution μ captures the optimism of agents about the state of the economy, while γ represents the precision of their beliefs about the fundamental. Precision γ is inversely related to the amount of uncertainty. As γ increases, the variance of beliefs decreases, and uncertainty declines.

Firms start each period with beliefs (μ, γ) and use all the information available to update their beliefs. By the end of the period, they have observed the public signals X and Z. Therefore, using Bayes' rule, the beliefs about next period's fundamental θ' are normally distributed with mean and

for the behavior of Italian industrial firms. Bikhchandani et al. (1998) survey the empirical social learning literature.

[9]We assume that the precision of each individual signal x_j is inversely proportional to $\overline{N}$ to prevent the signals to be fully revealing when we take the limit $\overline{N} \to \infty$, while preserving the positive relationship between economic activity and the amount of information. This captures the idea that uncertainty may subsist even when the number of firms is large, either because their information is correlated and arises from the same sources, or because large economies are more complex and subject to more shocks, preventing the learning problem from becoming trivial.

precision equal to

$$\mu' = \rho_\theta \frac{\gamma\mu + \gamma_z Z + n\gamma_x X}{\gamma + \gamma_z + n\gamma_x}, \tag{4}$$

$$\gamma' = \left(\frac{\rho_\theta^2}{\gamma + \gamma_z + n\gamma_x} + \left(1 - \rho_\theta^2\right) \sigma_\theta^2 \right)^{-1} \equiv \Gamma\left(n, \gamma\right). \tag{5}$$

These standard updating rules have straightforward interpretations: the mean of future beliefs μ' is a precision-weighted average of the present belief μ and the new signals, X and Z, whereas γ' depends on the precision of current beliefs, the precision of the signals, and the variance of the shock to θ. Importantly, the precision of future beliefs does not depend on the realization of the public signals, but only on n and γ. The higher is n, the more precise is the public signal X, and the lower is uncertainty in the next period. We define $\Gamma\left(n, \gamma\right)$ in (5) as the law of motion of the precision of information.

2.4 Firm Problem

We now describe the problem of a firm. In each period, given its individual fixed cost f and the common beliefs about the fundamental, a firm can either wait or invest. It solves the Bellman equation

$$V\left(\mu, \gamma, f\right) = \max\left\{V^W\left(\mu, \gamma\right), V^I\left(\mu, \gamma\right) - f\right\}, \tag{6}$$

where $V^W\left(\mu, \gamma\right)$ is the value of waiting and $V^I\left(\mu, \gamma\right)$ is the value of investing after incurring the investment cost f.

If a firm waits, it starts the next period with updated beliefs $\left(\mu', \gamma'\right)$ about the fundamental and a new draw of the fixed cost f'. Therefore, the value of waiting is

$$V^W\left(\mu, \gamma\right) = \beta \mathbb{E}\left[\int V\left(\mu', \gamma', f'\right) dF\left(f'\right) \mid \mu, \gamma\right]. \tag{7}$$

In turn, upon investment, a firm receives output θ and exits the economy. Therefore,

$$V^I\left(\mu, \gamma\right) = \mathbb{E}\left[u\left(\theta\right) \mid \mu, \gamma\right] = \frac{1}{a}\left(1 - e^{-a\mu + \frac{a^2}{2\gamma}}\right). \tag{8}$$

The firm's optimal investment decision takes the form of a cutoff rule $f^c\left(\mu, \gamma\right)$ such that a firm invests if and only if $f \leq f^c\left(\mu, \gamma\right)$. The cutoff is defined by the following indifference condition

$$f^c\left(\mu, \gamma\right) = V^I\left(\mu, \gamma\right) - V^W\left(\mu, \gamma\right). \tag{9}$$

2.5 Law of Motion for the Number of Investing Firms N

We now aggregate the individual decisions of the firms. As the investment decision follows the cutoff rule $f^c(\mu,\gamma)$, the process for the number of investing firms N satisfies

$$N\left(\mu,\gamma,\{f_j\}_{1\le j\le\overline{N}}\right)=\sum_{j=1}^{\overline{N}}\mathbb{I}\left(f_j\le f^c(\mu,\gamma)\right). \tag{10}$$

Since investment depends on a random fixed cost, the number of investing firms is a random variable that depends on the realization of the shocks $\{f_j\}_{1\le j\le\overline{N}}$. As these costs are i.i.d., the ex-ante probability of investment is identical across firms and equal to $F(f^c(\mu,\gamma))$. Therefore, the ex-ante distribution of N, as perceived by firms, is binomial,

$$N\mid\mu,\gamma\sim\text{Bin}\left(\overline{N},F\left(f^c(\mu,\gamma)\right)\right). \tag{11}$$

Note that N is only a function of the beliefs (μ,γ) and the individual shocks $\{f_j\}_{1\le j\le\overline{N}}$. Since these shocks are independent from the fundamental θ and the investment decisions are made before the observation of returns, there is nothing to learn from the non-investment of firms, nor from the realization of N itself.

2.6 Recursive Competitive Equilibrium

Focusing on the limiting case when $\overline{N}\to\infty$, the fraction of investing firms n becomes deterministic,

$$n=\frac{N}{\overline{N}}\xrightarrow{a.s}F\left(f^c(\mu,\gamma)\right).$$

We define a recursive competitive equilibrium as follows.[10]

Definition 1. A recursive competitive equilibrium consists of a cutoff rule $f^c(\mu,\gamma)$, value functions $V(\mu,\gamma,f)$, $V^W(\mu,\gamma)$, $V^I(\mu,\gamma)$, laws of motions for aggregate beliefs $\{\mu',\gamma'\}$, and a fraction of investing firms $n(\mu,\gamma)$, such that

1. The value function $V(\mu,\gamma,f)$ solves (6), with $V^W(\mu,\gamma)$ and $V^I(\mu,\gamma)$ defined according to (7) and (8), yielding the cutoff rule $f^c(\mu,\gamma)$ in (9);
2. The aggregate beliefs (μ,γ) evolve according to (4) and (5), under the perceived fraction of investing firms $n(\mu,\gamma)=F(f^c(\mu,\gamma))$.

[10] Fluctuations in N due to finite sampling are irrelevant for our purpose. Our results nonetheless carry on to the finite $\overline{N}$ case. Our existence and uniqueness proof for that case is available upon request.

3 Equilibrium Characterization

We first characterize the evolution of beliefs. We then show the existence and uniqueness of an equilibrium, and provide conditions under which firms are less likely to invest when uncertainty is high.

3.1 Evolution of Beliefs

The optimal investment rule $f^c(\mu, \gamma)$ depends on how beliefs evolve. We begin by establishing two simple lemmas about the dynamics of aggregate beliefs.

3.1.1 Evolution of the Mean of Beliefs

Using (4), we can characterize the stochastic process for the mean of beliefs as follows.

Lemma 1. *For a given n, the mean of beliefs μ follows an autoregressive process with time-varying volatility s,*

$$\mu^{'} = \rho_\theta \mu + s(n, \gamma)\, \varepsilon,$$

where $s(n, \gamma) = \rho_\theta \left(\frac{1}{\gamma} - \frac{1}{\gamma + \gamma_y + n\gamma_x} \right)^{\frac{1}{2}}$ *and* $\varepsilon \sim \mathcal{N}(0, 1)$.

The mean of beliefs captures the optimism of agents about the fundamental and evolves stochastically due to the the arrival of new information. It inherits the autoregressive property of the fundamental, and its volatility $s(n, \gamma)$ is time-varying because the amount of information that firms collect over time is endogenous. The volatility is decreasing with γ and increasing with n. In times of low uncertainty (γ high) agents place more weight on their current information and less on new signals, making the mean of beliefs more stable. In contrast, in times of high activity (n high) more information is released, making beliefs more likely to fluctuate.

3.1.2 Evolution of Uncertainty

The precision of beliefs γ reflects the inverse of uncertainty about the fundamental and its dynamics play a key role for the existence of uncertainty traps. Its law of motion satisfies the following properties.

Lemma 2. *The law of motion $\Gamma(n, \gamma)$ increases with n and γ. For a given fraction of investing firms n, the law of motion for the precision of beliefs $\gamma' = \Gamma(n, \gamma)$ admits a unique stable stationary point in γ.*

The thin solid curves on Figure 3 depict $\Gamma(n, \gamma)$ for different constant values of n. An increase in the level of activity raises the next period precision of information γ' for each level of γ in the current period. Since n is between 0 and 1, the support of the ergodic distribution of γ must lie between the bounds $\underline{\gamma}$ and $\overline{\gamma}$ defined by $\underline{\gamma} \equiv \Gamma(0, \underline{\gamma})$ and $\bar{\gamma} \equiv \Gamma(1, \overline{\gamma})$. In other words, $\underline{\gamma}$ is the stationary level of precision when no firm invests, while $\overline{\gamma}$ is the one when all firms invest.

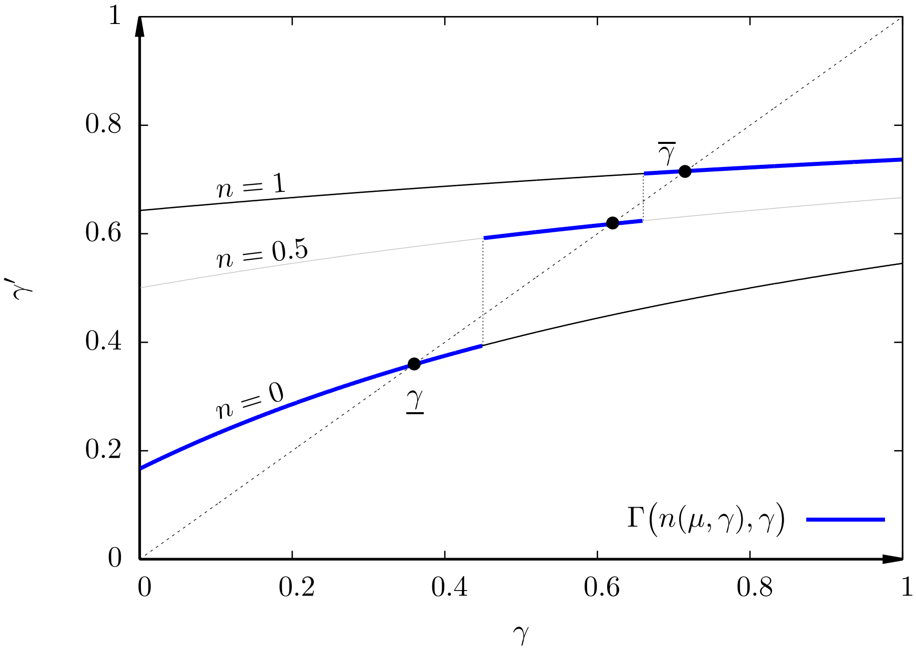


Figure 3: Example of dynamics for beliefs precision γ

In equilibrium, n varies with μ and γ. Suppose, as an example, that n is an increasing step function of γ that takes the values 0, 0.5 or 1, and let us keep μ fixed for the moment. Figure 3 illustrates how the feedback from uncertainty to investment opens up the possibility of multiple stationary points in the dynamics of the precision of beliefs, and therefore uncertainty. In this example, the function $\gamma' = \Gamma\left(n\left(\mu,\gamma\right),\gamma\right)$, depicted by the solid curve, has three fixed points. We formally establish, in Section 4, that this type of multiplicity can happen in equilibrium.

3.2 Existence and Uniqueness

We have described in Lemmas 1 and 2 how beliefs depend on the fraction of investing firms. We now characterize the equilibrium decision rule and provide existence and uniqueness conditions.

Proposition 1. *Under Assumptions 1-3, stated in Appendix F, and for γ_x sufficiently small, the equilibrium exists and is unique. Under some additional conditions satisfied when γ_x is small and risk aversion "a" is large enough, the equilibrium cutoff f^c is increasing in μ and γ.*

This proposition establishes the monotonicity of the equilibrium cutoff rule. Anticipating higher returns, a more optimistic firm (higher μ) is more likely to invest. In turn, uncertainty (lower γ) reduces the incentives to invest for two reasons. First, risk averse firms dislike uncertain payoffs. Second, since investment is costly and irreversible, there is an option value of waiting: in the face of uncertainty, firms prefer to delay investment to gather additional information and avoid downside risk.

It is essential for our mechanism that uncertainty discourages investment, a feature typical of optimal stopping time models of investment. Assumption 3, satisfied if the persistence of the fundamental is high enough and its volatility is sufficiently low, ensures that the fundamental does not vary too much over time, so that firms have an incentive to wait in order to collect more information.[11] This condition alone, however, is not sufficient in our context. The monotonicity of the cutoff f^c in γ requires additional restrictions because of the endogeneity of beliefs, which gives rise to ambiguous feedback effects. For instance, the variation in n implied by fluctuations in μ or γ affect the volatility of next period's mean beliefs μ' (Lemma 1). This, in turn, can have ambiguous effects on firms' current incentives to invest. To ensure that the first-order effects of risk aversion and option value dominate, we must bound these feedback effects. Since they operate solely through social learning, we can do so by imposing an upper bound on the informativeness of this channel, γ_x. When γ_x is small enough, the equilibrium cutoff is guaranteed to be increasing in μ and γ, as one would expect in the absence of social learning.

Establishing the existence of an equilibrium is relatively straightforward, because the problem is continuous and general fixed point theorems apply. Showing uniqueness is more challenging because our economy features complementarities in information: the more firms invest, the more uncertainty declines, encouraging further investment. If these complementarities are strong enough, they can lead to multiple equilibria. To prevent this, we use again the insight that the magnitude of this feedback is governed by the precision γ_x of the social learning channel. We show, in particular, that the main fixed point problem that characterizes the optimal cutoff rule is a contraction, and that it is therefore unique, when γ_x is small.[12] The uniqueness of the equilibrium is an attractive feature, as it leads to unambiguous predictions and makes the model amenable to quantitative work. Despite the uniqueness of the equilibrium, the model features interesting non-linear dynamics and multiple stationary points, as we show in Section 4.

Figure 4 illustrates how the investment probability varies as a function of beliefs (μ, γ) when monotonicity obtains. The fraction of investing firms increases as they are more optimistic (μ high) or less uncertain (γ high) about the fundamental.

[11] The law of motion (5) highlights the importance of the persistence of the fundamental ρ_θ for the dynamics of uncertainty. As ρ_θ declines, past observations contain less information about the current value of the fundamental and learning therefore becomes less relevant. At a result, the option value of waiting becomes smaller and the conditions for uncertainty traps to exist, provided in the next section, are less likely to be satisfied.

[12] We show that the mapping that characterizes the optimal cutoff rule is a contraction in the space of Lipschitz continuous functions for some given moduli, which allows us to put a bound on these feedback effects. We cannot rule out the existence of equilibrium cutoffs that do not satisfy this property. We can, however, explicitly rule them out in the case of the planner's allocation, where Lipschitz continuity is necessarily satisfied.

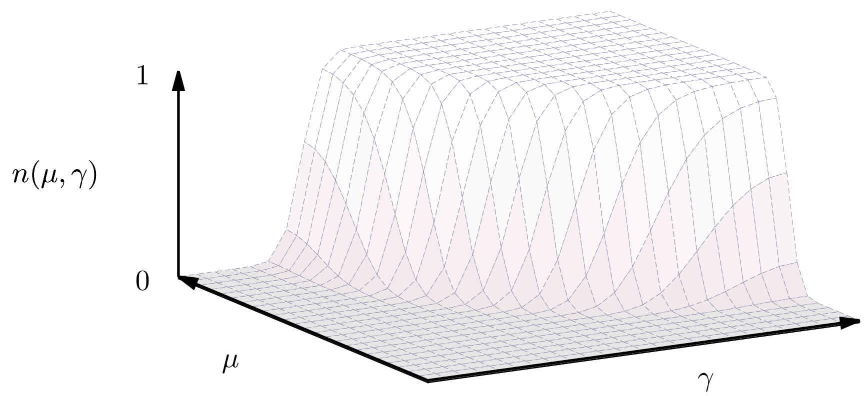


Figure 4: Fraction of investing firms $n(\mu,\gamma)$

4 Uncertainty Traps

We now examine the interaction between firms' behavior in the face of uncertainty and social learning. This interaction leads to episodes of self-sustaining uncertainty and low activity, which we call uncertainty traps. We provide sufficient conditions on the parameters that guarantee the existence of such traps and discuss the type of aggregate dynamics that they imply. We find that the response of the economy to shocks is highly non-linear: it quickly recovers after small shocks, but large, short-lived shocks may plunge the economy into long-lasting recessions. We also characterize the constrained planner's problem and discuss its policy implications.

4.1 Definition and Existence

We define *uncertainty traps* as the coexistence of multiple stationary points in the dynamics of belief precision — a situation similar to the one depicted in Figure 3.

Definition 2. There is an *uncertainty trap* if there exists an interval (μ_l,μ_h) such that, for every $\mu \in (\mu_l,\mu_h)$, there are at least two locally stable fixed points in the dynamics of the precision of beliefs $\gamma' = \Gamma(n(\mu,\gamma),\gamma)$.

We refer to these multiple stationary points as *regimes.* Note that multiplicity of regimes does not imply multiplicity of equilibria. This distinction is important because it highlights that the model is not subject to indeterminacy. While multiple values of γ may satisfy the equation $\gamma = \Gamma(n(\mu,\gamma),\gamma)$ for a given μ, the regime that prevails at any given time is unambiguously determined by the history of past aggregate shocks, summarized by the current beliefs (μ,γ). The definition of uncertainty traps also emphasizes the notion of *stability,* which is required for the type of self-sustaining dynamics that we describe. Notice, however, that we only require local stability along the dimension γ while μ keeps evolving according to its law of motion.

The following proposition formally establishes that uncertainty traps exist for a range of mean of beliefs μ under some condition on the dispersion of investment costs.

Proposition 2. *Under the conditions of Proposition 1 and one additional condition satisfied for σ_f small enough or risk aversion "a" high enough, the economy features an uncertainty trap with at least two regimes $\gamma_l(\mu) < \gamma_h(\mu)$ for $\mu \in (\mu_l, \mu_h)$. Regime γ_l is characterized by high uncertainty and low investment, while regime γ_h is characterized by low uncertainty and high investment.*

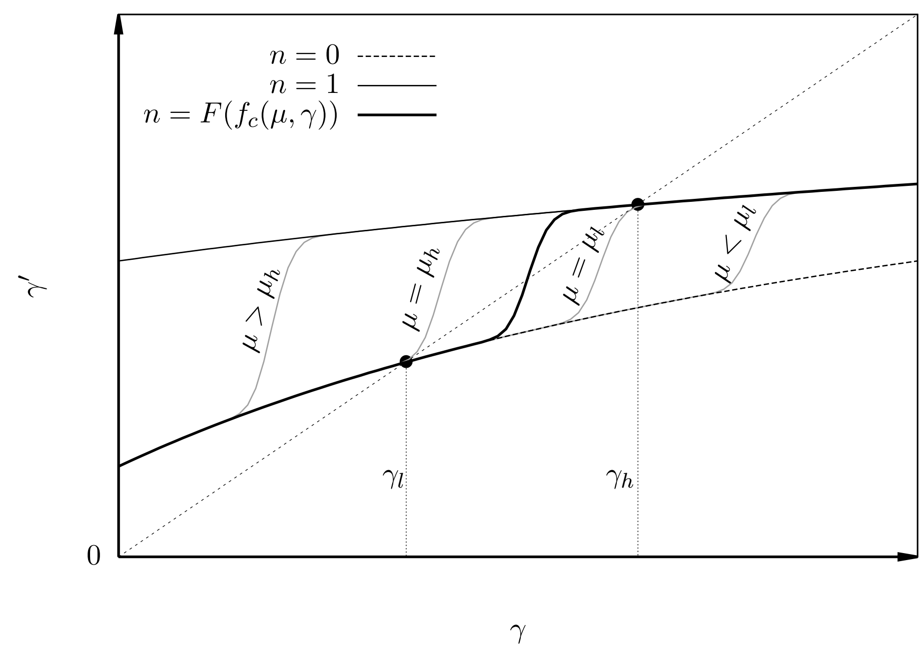


Figure 5: Dynamics of the precision of beliefs $\gamma' = \Gamma(n(\mu, \gamma), \gamma)$ for different values of μ

Figure 5 presents examples for the law of motion of γ when the investment costs f are normally distributed. The solid curves represent the function $\gamma' = \Gamma(n(\mu, \gamma), \gamma)$ evaluated at five different values of μ, with the thick solid curve corresponding to an intermediate value of μ. In all cases, for small γ, uncertainty is high and firms do not invest. As a result, they do not learn from observing economic activity and the precision of beliefs γ' remains low. As the precision γ increases, uncertainty decreases and firms become sufficiently confident about the fundamental to start investing. As that happens, uncertainty decreases further.

In our example, the thick curve intersects the 45° line three times. The second intersection corresponds to an unstable regime, but the other two are locally stable. We denote these regimes by γ_l and γ_h. In regime γ_l, uncertainty is high and investment is low, while the opposite is true in regime γ_h.

Proposition 2 shows that this situation is a generic feature of the equilibrium when the dispersion of investment costs σ_f is small. This condition ensures that the feedback of investment on information is strong enough to sustain distinct stationary points.

4.2 Dynamics: Non-linearity and Persistence

We now describe the full dynamics of the economy by taking into account the evolution of μ in response to the arrival of new information. Figure 5 shows that, as long as μ stays between the values μ_l and μ_h, defined in Proposition 2, the two regimes $\gamma_l(\mu)$ and $\gamma_h(\mu)$ preserve their stability. As a result, uncertainty and the fraction of active firms n are relatively unaffected by changes in μ. In contrast, for values of μ above μ_h, a large enough fraction of firms invest, so the dynamics of beliefs only admits the high-activity regime as a stationary point. Similarly, for values below μ_l, the economy only admits the low-activity regime. Therefore, sufficiently large shocks to μ can make one regime disappear and trigger a regime switch.

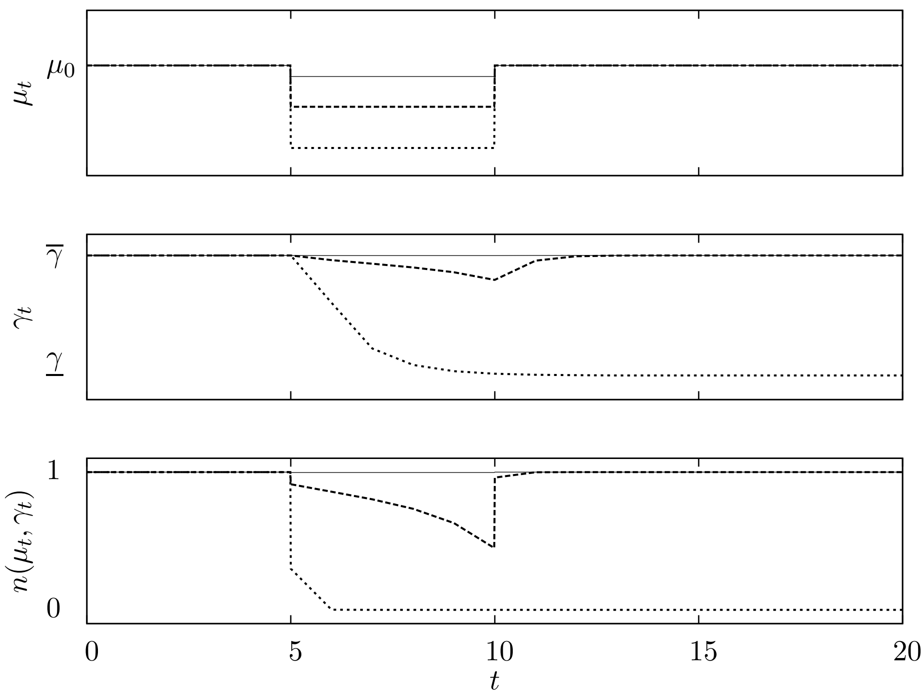


Figure 6: Persistent effects of temporary shocks

The economy displays non-linear dynamics: it reacts very differently to large shocks in comparison to small ones. Figure 6 shows various simulations to illustrate this feature using the example from Figure 5. The top panel presents three different series of shocks to the mean of beliefs μ. The three series start from the high-activity/low-uncertainty regime. At $t = 5$, the economy is hit by a negative shock to μ, due to a bad realization of either the public signals or the fundamental. The mean of beliefs then returns to its initial value at $t = 10$. Across the three series, the magnitude of the shock is different.

The middle and bottom panels show the response of beliefs precision γ and the fraction of investing firms n. The solid gray line represents a small temporary shock, such that μ remains within (μ_l, μ_h). Despite the negative shocks to the mean of beliefs, all firms keep investing and the precision of beliefs is unaffected. When the economy is hit by a temporary shock of medium size (dashed line), some firms stop investing, leading to a gradual increase in uncertainty. As uncertainty rises, investment falls further and the economy starts to drift towards the low regime. However, when the mean of beliefs recovers, the precision of information and the number of active firms quickly return to the high-activity regime. In contrast, when the economy is hit by a large temporary shock (dotted line), the number of firms delaying investment is large enough to produce a self-sustaining increase in uncertainty. The economy quickly shifts to the low-activity regime and remains there even after the mean of beliefs recovers.

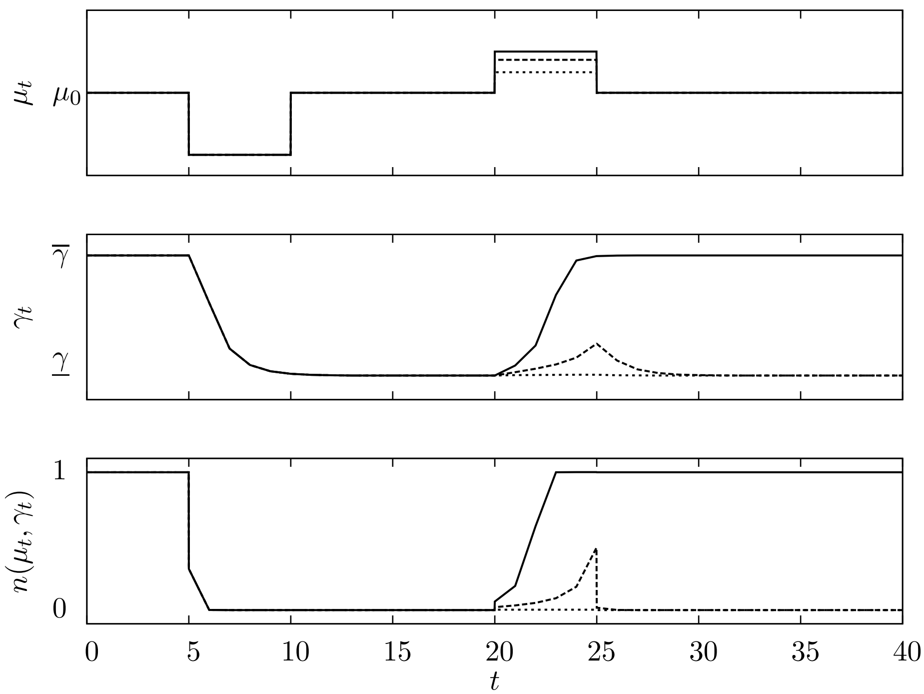


Figure 7: Escaping an uncertainty trap

We now discuss how the economy escapes from the trap in which it fell in Figure 6. Figure 7 shows the effect of positive shocks when the economy starts from the low regime. The economy receives positive signals that lead to a temporary increase in mean beliefs between periods 20 and 25, possibly because of a recovery in the fundamental. When the temporary increase in average beliefs is not sufficiently strong, the recovery is interrupted as μ returns to its initial value. However, when the temporary increase is sufficiently large, the economy reverts back to the high-activity regime. Once again, temporary shocks of sufficient magnitude to the fundamental may lead to nearly permanent effects on the economy.

4.3 Additional Remarks

A number of additional lessons can be drawn from these simulations. First, in our framework, uncertainty is a by-product of recessions. This result echoes the empirical findings of

Bachmann et al. (2013) who show that uncertainty is partly caused by recessions and conclude, by that, that it is of secondary importance for the business cycle. We show, however, that uncertainty may still have a large impact on the economy by affecting the persistence and depth of recessions, even if it is not what triggers them.

Second, as in models with learning in the spirit of Van Nieuwerburgh and Veldkamp (2006), our theory provides an explanation for asymmetries in business cycles. In good times, since agents receive a large flow of information, they react faster to shocks than in bad times.

Third, our economy may feature high uncertainty without volatility. For instance, in the low regime, agents are highly uncertain about the fundamental but the volatility of economic aggregates is low. Therefore, according to our theory, subjective uncertainty may affect economic fluctuations even if no volatility is observed in the data. This distinguishes our approach from the existing uncertainty-driven business cycle literature in the spirit of Bloom (2009). In particular, direct measures of subjective uncertainty rather than measures of volatility are important to capture the full amount of uncertainty in the economy.

Finally, a recent literature (Bachmann et al., 2013; Orlik and Veldkamp, 2013) uses survey data to derive measures of uncertainty based on ex-ante forecast errors. Our model highlights a potential shortcoming of this approach, as uncertainty about *fundamentals* differs from uncertainty about *endogenous variables*, such as output or investment. For example, when the economy is trapped in the low activity regime, firms know that all firms are uncertain, and therefore that investment is likely to be low, such that the economy is less exposed to aggregate risk. As a result, their forecasts about economic aggregates are accurate, even though their uncertainty about the fundamental is high. As implied by the model, forecast errors about variables like output may not always be a good proxy for uncertainty about fundamentals.

4.4 Policy Implications

The economy is subject to an information externality: in the decentralized equilibrium, firms invest less often than they should because they do not internalize the release of information to the rest of the economy caused by their investment. In Proposition 3, we solve the problem of a constrained planner subject to the same information technology as individual agents. We show that the decentralized economy is constrained inefficient, and that an investment subsidy is sufficient to restore constrained efficiency.

Proposition 3. *Under Assumptions 1-3, stated in the Appendix, the recursive competitive equilibrium is constrained inefficient. The efficient allocation can be implemented with positive investment subsidies* $\tau(\mu, \gamma)$ *and a uniform tax.*

The subsidy that implements the optimal allocation takes a simple form to align social and private incentives. As shown in the proof of the proposition, it is simply the sum of the social value of releasing an additional signal to the economy and the private value of delaying investment.

The optimal policy being a subsidy, Proposition 3 implies that firms are more likely to invest in the efficient allocation than in the laissez-faire economy. However, uncertainty traps can still arise in the efficient allocation. Proposition 4 below establishes the result.

Proposition 4. *Under Assumptions 1-2 and γ_x small enough, the planner's allocation is subject to an uncertainty trap for σ_f low enough and risk aversion "a" high enough.*

The existence of uncertainty traps in the planner's allocation may be surprising if one thinks of the planner as a coordinator that should always prefer the high regime, as one might expect in a model with multiple equilibria. As it turns out, transitioning from one regime to the other is costly and risky. If the planner does not have more information than individual agents, it is still optimal to wait when uncertainty is high enough. Hence, there may still exists a sufficiently strong feedback from beliefs to actions in the constrained-efficient allocation to generate uncertainty traps. However, while uncertainty traps remain present in the efficient allocation, they are less likely to arise than in the laissez-faire economy because firms have stronger incentives to invest.

5 Quantitative Evaluation

To evaluate the quantitative importance of uncertainty traps, we now embed the mechanism into a general equilibrium macroeconomic framework. We first describe the quantitative framework and its parametrization. We then compare the model's implications for standard business cycle moments with the data and two alternative models: the RBC model and a restricted version of our model in which uncertainty about the fundamental is fixed over time. Finally, we present our main quantitative exercise, in which we compare the behavior of economic aggregates in the data against the predictions from our model and alternative models in the context of the Great Recession.

5.1 Quantitative Model with Uncertainty Traps

We extend the baseline model along several dimensions. First, firms are now long-lived, use both capital and labor to produce, and accumulate capital over time. They enter the economy endogenously depending on economic conditions and exit exogenously. Second, firms are owned by a risk-averse representative household that maximizes utility over consumption and leisure. Third, factor and goods prices are endogenously determined in general equilibrium. As in the baseline model, firms must pay an irreversible fixed cost to operate and social learning takes place when a firm begins to produce. As a result, the number of entering firms responds to uncertainty about the fundamental, and uncertainty depends on economic activity.

5.1.1 Preferences and Technology

The representative household chooses consumption C_t and labor L_t to maximize the expected discounted sum of future utility

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^t U\left(C_t, L_t\right), \tag{12}$$

where $0 < \beta < 1$ is the discount rate. The household supplies labor in a perfectly competitive market at a wage w_t. It also owns the firms in the form of claims to their dividends.

A single good used for consumption and investment is produced by a continuum of firms of measure m. Each firm $j \in [0, m]$ produces the final good by operating a Cobb-Douglas technology,

$$A\left(1+\theta\right)\left(k_j^{\alpha} l_j^{1-\alpha}\right)^{\omega},\ 0 < \alpha < 1,\ 0 < \omega < 1,$$

using l_j units of labor and k_j units of capital. The parameter $0 < \alpha < 1$ controls the capital intensity. The firm-level returns to scale, or span-of-control (Lucas Jr, 1978), parameter ω is assumed to be strictly less than one to deliver a well-defined notion of firm size. The fundamental θ follows the AR(1) process $\theta' = \rho_\theta \theta + \varepsilon^\theta$, with $\varepsilon^\theta \sim$ iid $\mathcal{N}\left(0, \left(1-\rho_\theta^2\right)\sigma_\theta^2\right)$.[13] The total mass of firms m evolves endogenously. Each period, a mass $Q > 0$ of potential entrants has the option to start production, but only an (endogenous) fraction n of them does so. The mass Q remains fixed over time. Firms exit at an exogenous rate $\delta_m > 0$.

Each period, firms pay a fixed cost f common across firms and denominated in units of the final good. We assume that $f \sim \mathcal{N}\left(\mu_f, \sigma_f^2\right)$ is drawn independently over time.[14] Due to the irreversibilities created by these fixed costs, fewer firms enter in times of heightened uncertainty.

5.1.2 Information and Timing

As in the baseline model, agents do not observe the true value of the fundamental θ but learn about it from two sources. First, they learn from a public signal Z. In contrast to the baseline model, where this signal captured exogenous information released by media and statistical agencies, we now explicitly model the signal Z as a summary of the information collected through the observation of certain economic aggregates. As in any model with information frictions, restrictions about what agents observe must be imposed to avoid perfectly revealing the fundamental. Agents cannot, for instance, perfectly observe output as this would reveal θ. We assume instead that agents are able to observe the value added of each firm, as well as its aggregate counterpart, but are unable to perfectly distinguish between its individual components: revenue and fixed cost.[15] As a result, a

[13] The additive specification of TFP, $1+\theta$, ensures that the variance of beliefs about θ does not affect expected output directly. As in our calibration the standard deviation of the ergodic distribution of θ is much smaller than 1, productivity is always positive in our simulations.

[14] The baseline theory included an idiosyncratic component to these fixed costs, which we ignore here for simplicity. Appendix E.4 performs sensitivity analysis on this assumption. We assume, however, that f is subject to aggregate shocks to be consistent with our information structure, as we explain in the next subsection.

[15] This assumption is in the spirit of Lucas (1972), where firms cannot distinguish between real and nominal shocks. A previous version of the paper assumed that firms cannot distinguish between aggregate and idiosyncratic

high level of value added may reflect either a high value of the fundamental θ or a low value of the fixed costs.[16] Second, and more specific to the channel we study in this paper, agents also learn from signals emanating from others. As in the $\overline{N} \to \infty$ case of the baseline model, the entry of an infinitesimal measure of firms dj releases a normally distributed signal x_j about θ, observed by everyone, with a precision $\gamma_x dj$, proportional to the mass of entrants. Again, the information collected through this social learning channel can be summarized by a public signal X with precision $nQ\gamma_x$.[17] All signals being public, beliefs are common across firms and the representative household.

In each period, events unfold as follows:

1. Incumbent firms, potential entrants and the household start with the same prior distribution over the fundamental, $\theta \mid \mathcal{I} \sim \mathcal{N}\left(\mu, \gamma^{-1}\right)$. The fundamental θ and the fixed cost f are drawn but unobserved.

2. The Q potential entrants decide whether to enter or not. A fraction n of them enters and start producing next period.

3. The m incumbent firms choose labor and investment. The household decides how much labor to supply and the labor market clears.

4. Fixed costs are paid, and production takes place. All agents observe the signal Z, which captures the information contained in value added, and the signal X from new entrants, and update their beliefs. A fraction δ_m of firms exogenously exits.

5.1.3 Firm-Level Problem

The aggregate state space of the economy is (μ, γ, K, m) where $K = \int_0^m k_j dj$ is the aggregate capital stock. Realized individual profits for a firm operating with k units of capital and l units of labor are[18]

$$\pi\left(k, l; \mu, \gamma, K, m, \theta, f\right) = A\left(1+\theta\right) k^{\alpha\omega} l^{(1-\alpha)\omega} - w\left(\mu, \gamma, K, m\right) l - f. \tag{13}$$

The value of an incumbent firm that has accumulated k units of capital is then

$$\begin{aligned} V^I\left(k; \mu, \gamma, K, m\right) = \max_{k', l} \mathbb{E}\Big\{ & U_c\left(C, L\right)\left[\pi\left(k, l; \mu, \gamma, K, m, \theta, f\right) + \left(1-\delta_K\right) k - k'\right] \\ & + \delta_m U_c\left(C, L\right) k' + \beta\left(1-\delta_m\right) V^I\left(k'; \mu', \gamma', K', m'\right) \mid \mu, \gamma \Big\}, \end{aligned} \tag{14}$$

productivity shocks. The benefit of our current approach is to allow for a simple aggregation of the economy.

[16] Despite this restriction on the observability of gross output and θ, all other economic aggregates are observed and agents use all the available information to make their decision. However, as the timing will make clear, observing other variables such as the wage rate w, consumption C, the aggregate capital stock K, aggregate employment L, aggregate value added, the measure of entrants n, or the measure of incumbents m does not reveal any additional information.

[17] A formal derivation of this information aggregation result is in Subsection 5.1.5.

[18] Note that the presence of fixed operating costs could lead to negative profits. Since f is small relative to output in our calibration, this virtually never happens.

subject to the laws of motion for the aggregate state variables $\{\mu, \gamma, K, m\}$, described in the following sections. The parameter $0 < \delta_K < 1$ is the depreciation rate of capital. This firm chooses labor l and next-period capital k' to maximize the expected sum of profits, discounted by the marginal utility $U_c(C, L)$, which plays the role of the stochastic discount factor. When a firm exits, which happens with probability δ_m, its accumulated capital is scrapped and returned to the household at the end of the period.

Consider now the problem of a potential entrant. In each period, a potential entrant decides between waiting and entering. Its value is

$$V(\mu, \gamma, K, m) = \max \left\{V^W(\mu, \gamma, K, m), V^E(\mu, \gamma, K, m)\right\}. \tag{15}$$

If a potential entrant waits, it preserves the option of entering next period; hence the value of waiting is

$$V^W(\mu, \gamma, K, m) = \beta \mathbb{E}\left[V\left(\mu', \gamma', K', m'\right) | \mu, \gamma\right]. \tag{16}$$

If, instead, the potential entrant decides to enter, its value is

$$V^E(\mu, \gamma, K, m) = \max_{k'_e} (1 - \delta_m) \mathbb{E}\left[-U_c(C, L) k'_e + \beta V^I\left(k'_e; \mu', \gamma', K', m'\right) | \mu, \gamma\right]. \tag{17}$$

The definition of V^E indicates that a potential entrant chooses the amount of capital k'_e, carried into the next period if it survives the δ_m shock or returned to the household at the end of the period otherwise. Upon entry, its value next period is equal to the value of an incumbent firm that has accumulated k'_e units of capital, $V^I(k'_e; \mu', \gamma', K', m')$.

5.1.4 Aggregates

Incumbent and entrants face the same investment problem and choose the same next-period capital level $k'(\mu, \gamma, K, m)$. Therefore, next-period's aggregate capital stock is

$$K'(\mu, \gamma, K, m) = m'(\mu, \gamma, K, m) k'(\mu, \gamma, K, m), \tag{18}$$

where $m'(\mu, \gamma, K, m)$ is the mass of incumbents next period, given by

$$m'(\mu, \gamma, K, m) = (1 - \delta_m)(m + n(\mu, \gamma, K, m) Q). \tag{19}$$

The fraction of entering firms among the Q potential entrants $n(\mu, \gamma, K, m)$ must be consistent with individual entry decisions, in the sense that

$$n(\mu, \gamma, K, m) = \begin{cases} 1 & \text{if } V^E(\mu, \gamma, K, m) > V^W(\mu, \gamma, K, m) \\ \in [0, 1] & \text{if } V^E(\mu, \gamma, K, m) = V^W(\mu, \gamma, K, m) \\ 0 & \text{if } V^E(\mu, \gamma, K, m) < V^W(\mu, \gamma, K, m). \end{cases} \tag{20}$$

In turn, aggregate labor demand is

$$L(\mu, \gamma, K, m) = m \times l\left(\frac{K}{m}; \mu, \gamma, K, m\right), \tag{21}$$

where $l(k; \mu, \gamma, K, m)$ is the firm-level labor demand resulting from (14).

5.1.5 Information and Beliefs

We now characterize the information contained in the signals observed by the agents. First, as in the baseline model, the information diffused through social learning can be aggregated into a single signal X which averages the individual signals released by entrants,[19]

$$X = \frac{1}{nQ}\int_0^{nQ} x_j dj = \theta + \varepsilon^X,\ \varepsilon^X \sim \mathcal{N}\left(0, (nQ\gamma_x)^{-1}\right), \tag{22}$$

where $nQ\gamma_x$, the endogenous precision of the social learning channel, changes with economic activity. Second, the information conveyed by observing value added is equivalent to the information conveyed by the signal $Z \sim \mathcal{N}\left(\theta, (\gamma_z(K, L, m))^{-1}\right)$ with precision[20]

$$\gamma_z(K, L, m) = \left[A\left(\frac{K^\alpha L^{(1-\alpha)}}{m}\right)^\omega \frac{1}{\sigma_f}\right]^2.$$

In contrast to the benchmark model, the precision γ_z of this signal now changes with economic activity — a natural implication of assuming that agents observe economic aggregates instead of the fundamental directly. In our calibrated economy, we find that fluctuations in γ_z, which depend on the *stock* m of incumbent firms, are considerably smaller than fluctuations in the precision of X, which depends on the *flow* n of incumbent firms. Therefore, the endogenous uncertainty in the economy largely evolves as a function of the X signal.[21]

As in the baseline model, agents are fully rational and use all information available to update their beliefs according to Bayes' Law. The laws of motion for the mean and the precision of beliefs

[19] As in the infinite $\overline{N}$ case in the baseline model, X in expression (22) is to be understood as the distributional limit of the average of N signals with precisions $\gamma_x/\overline{N}$, i.e., $\lim \frac{1}{N}\sum_1^N x_j \sim \mathcal{N}\left(0, \left(\gamma_x N/\overline{N}\right)^{-1}\right)$ as $\overline{N} \to \infty$ and $N/\overline{N} \to nQ$.

[20] Since all incumbent firms are identical, individual value added is $A(1+\theta)\left(\frac{K}{m}\right)^{\alpha\omega}\left(\frac{L}{m}\right)^{(1-\alpha)\omega} - f$. Since K, L, m and the distribution of f are known, observing value added is equivalent to observing $\theta - \frac{m^\omega}{AK^{\alpha\omega}L^{(1-\alpha)\omega}}(f - \mu_f) \sim \mathcal{N}\left(\theta, \left[Am^{-\omega}K^{\alpha\omega}L^{(1-\alpha)\omega}\right]^{-2}\sigma_f^2\right)$.

[21] In our calibrated economy, fluctuations in γ_z only accounts for 2.7% of the total fluctuation in uncertainty while social learning through X accounts for the rest.

are

$$\mu' = \rho_\theta \frac{\gamma\mu + \gamma_z Z + nQ\gamma_x X}{\gamma + \gamma_z + nQ\gamma_x}, \tag{23}$$

$$\gamma' = \left(\frac{\rho_\theta^2}{\gamma + \gamma_z + nQ\gamma_x} + \left(1 - \rho_\theta^2\right) \sigma_\theta^2 \right)^{-1}. \tag{24}$$

5.1.6 Recursive Competitive Equilibrium

We are now ready to define a competitive equilibrium for this economy.

Definition 3. A recursive competitive equilibrium is a collection of value functions $V(\mu,\gamma,K,m)$, $V^W(\mu,\gamma,K,m)$, $V^E(\mu,\gamma,K,m)$ and $V^I(k;\mu,\gamma,K,m)$ individual policy functions $k'(\mu,\gamma,K,m)$, $k'_e(\mu,\gamma,K,m)$ and $l(k;\mu,\gamma,K,m)$, aggregate policy functions $K'(\mu,\gamma,K,m)$, $m'(\mu,\gamma,K,m)$, $\mu'(\mu,\gamma,K,m)$, $\gamma'(\mu,\gamma,K,m)$, $n(\mu,\gamma,K,m)$, $L(\mu,\gamma,K,m)$ and $C(\mu,\gamma,K,m,\theta,f)$, and wages $w(\mu,\gamma,K,m)$ such that

1. The value functions $V(\mu,\gamma,K,m)$, $V^W(\mu,\gamma,K,m)$, $V^E(\mu,\gamma,K,m)$ and $V^I(k;\mu,\gamma,K,m)$, and the associated policy functions $k'(\mu,\gamma,K,m)$, $k'_e(\mu,\gamma,K,m)$ and $l(k;\mu,\gamma,K,m)$, solve the Bellman equations (14)-(17) under the entry schedule $n(\mu,\gamma,K,m)$ and the laws of motion $K'(\mu,\gamma,K,m)$, $m'(\mu,\gamma,K,m)$, $\mu'(\mu,\gamma,K,m)$ and $\gamma'(\mu,\gamma,K,m)$ given by (18), (19), (23) and (24);
2. The fraction of entering firms $n(\mu,\gamma,K,m)$ satisfies the consistency equation (20);
3. The policy functions $L(\mu,\gamma,K,m)$ and $C(\mu,\gamma,K,m,\theta,f)$ solve the household's first order condition on labor supply,

$$\frac{\mathbb{E}\left[U_L\left(C(\mu,\gamma,K,m,\theta,f), L(\mu,\gamma,K,m)\right)\right]}{\mathbb{E}\left[U_C\left(C(\mu,\gamma,K,m,\theta,f), L(\mu,\gamma,K,m)\right)\right]} = w(\mu,\gamma,K,m);$$

4. The aggregate resource constraint is satisfied:

$$C(\mu,\gamma,K,m,\theta,f) + K'(\mu,\gamma,K,m) - (1-\delta_K)K + mf = A(1+\theta)m^{1-\omega}\left(K^\alpha L(\mu,\gamma,K,m)^{1-\alpha}\right)^\omega.$$

5.2 Calibration

5.2.1 Standard Parameters

The time period is one quarter. Most of the moments that we target are computed starting in 1978:Q1, when the Longitudinal Business Database (LBD) that we use for firm-level moments begins, and stopping in 2007:Q3, at the onset of the 2007-2009 recession, allowing us to evaluate the out-of-sample properties of the model in our Great Recession exercise. In our benchmark specification, we assume GHH preferences, $U = \log\left(C - L^{1+\nu}/(1+\nu)\right)$, and we also report the results of our main quantitative exercise under CRRA preferences in Appendix C as robustness.[22]

[22]GHH preferences are common in the information frictions literature. We adopt them in our benchmark specification because the usual CRRA preferences generate a counterfactual correlation between economic activity and

We set the Frisch elasticity $\nu = 2.84$ which corresponds to the average aggregate Frisch elasticity of hours reported by Chetty et al. (2011). The discount rate β is chosen to match an annual value of 0.95. The depreciation rate is set to an annual value of 0.1.

For the production function parameters, we normalize $A = 1$ and set the returns-to-scale parameter ω to 0.89, which corresponds to the weighted average across 2-digit SIC estimates of the returns to scale from Basu and Fernald (1997).[23] We set the capital intensity parameter α so that $(1-\alpha)\,\omega = 0.645$ to match the average labor compensation over GDP from 1978-2007 according to annual data from the Penn World Table (Feenstra et al., 2015).[24]

We set $\delta_m = 2.6\%$, which corresponds to the employment-weighted firm exit rate for all firms in the Longitudinal Business Database between 1978 and 2007. The mass of potential entrants Q is normalized to 1.[25]

The parameters $\left\{\rho_\theta, \sigma_\theta^2\right\}$ of the fundamental process θ are estimated using the quarterly utilization-adjusted TFP series from Fernald (2014) over 1978Q1-2007Q3 after removing a linear trend. This yields $\sigma_\theta = 0.028$ and $\rho_\theta = 0.964$.

5.2.2 Information and Fixed-Cost Distribution

With all the above parameters calibrated, it only remains to set values for the precision of individual signals γ_x and the mean and variance of the distribution of fixed costs $\left\{\mu_f, \sigma_f^2\right\}$. These parameters govern the option-value effects and the evolution of Bayesian uncertainty about TFP. To the best of our knowledge, no direct empirical measure exists for this concept of uncertainty. The variance of beliefs about θ, however, is tightly related in our model to the ex-ante forecast variance about endogenous variables like output. We thus target moments of the distribution of uncertainty in output growth forecast provided by the Survey of Professional Forecasters (SPF) and use this series as our main empirical proxy for uncertainty.

The SPF asks a panel of forecasters to provide the *distribution* of their beliefs about the growth rate of real output in percentage terms between the current year and the last, and these distributions are averaged across forecasters.[26] We compute the standard deviation of this averaged distribution in every year, and use moments of its time series to calibrate the model. To fit the parameters, we compute the exact same object in a long-run simulation of our model. We pick the values of $\{\gamma_x, \mu_f, \sigma_f\}$ by targeting the mean, the 5th percentile and the 95th percentile of the empirical distribution of uncertainty over the 1992-2007 period, corresponding to the time period over which

positive signals (see Beaudry and Portier (2006) for a VAR estimation of the impact of news shocks). Upon receiving a positive signal about the economy, the wealth effect on the labor supply leads to a decline in output on impact. See Jaimovich and Rebelo (2009) for a discussion of the role of preferences in the news-shocks literature.

[23] Our parametrization corresponds to the γ^v parameter for the private economy estimated using OLS from Table 2 in Basu and Fernald (1997).

[24] Specifically we use the series LABSHPUSA156NRUG from the FRED database.

[25] For any given Q, we can replicate the aggregate allocation, albeit with a different measure of firms m, by rescaling A, γ_x, μ_f and σ_f.

[26] Specifically, at the beginning of each quarter, the SPF asks each forecaster to report the probability that growth between the previous year and the current year will fall within each of several bins. We use forecast reported in the 4th quarter, which represents a measure of uncertainty over just one quarter, because it maps easily into our model.

data on the distribution of real GDP growth is available in the SPF up to the beginning of the Great Recession.

These three moments are directly informative about the three parameters that we need to calibrate. The 95th percentile corresponds to periods of high uncertainty about real output growth, which in the model correspond to periods with low firm entry when uncertainty is mostly driven by the aggregate public signal Z, as opposed to the social learning signal X. Therefore, the 95th percentile is useful to identify σ_f which governs the informativeness of Z. Similarly, the 5th percentile corresponds to periods of low uncertainty, which in the model correspond to periods of high firm entry, when γ_x is the main driver of uncertainty. The average fixed cost μ_f affects the average fraction of entrants n and relates to the average level of uncertainty.

The parameters are estimated by minimizing an equal-weighted distance between the empirical and simulated moments. The numerical algorithm used to solve the model is described in Appendix B. Table 1 reports the fit. The calibrated parameters are $\gamma_x = 450$, $\mu_f = 0.0115$ and $\sigma_f = 0.0155$. As the table shows, the calibrated model cannot match all moments at the same time. In particular, because TFP is the only source of uncertainty while the SPF forecasters may worry about other shocks, we have difficulty matching the upper tail of uncertainty in the data and there is on average less uncertainty in our model. We thus view our results as conservative on the role of uncertainty in the economy.

Uncertainty About GDP Growth	Data (%)	Model (%)
Mean	0.60	0.55
5th percentile	0.45	0.50
95th percentile	0.73	0.64

Notes. Uncertainty is computed as the standard deviation of the SPF distribution over 1992Q4 to 2007Q4 of current year's annual over last year's annual real GDP stated in the last quarter of the current year. Growth rates and standard deviations are stated in percentage terms. We use the 5% and 95% percentiles, instead of the min and the max, for robustness against outliers. Uncertainty in our model is computed over a simulation of 50,000 periods using the same definition as in the data.

Table 1: Calibrated Moments from the Survey of Professional Forecasters

5.3 Business-Cycle Moments

We now evaluate the performance of our model in explaining standard business cycle moments. For that, we compare the benchmark model that we have just described against the data and against two alternative models. The first alternative model ("RBC") is the standard real business cycle model under complete information, identically parametrized. The second model ("fixed θ-uncertainty") is a version of our model in which firms update their beliefs about the mean of the fundamental θ (i.e., they update μ using (23)) but uncertainty (i.e., the inverse of the precision of beliefs γ) remains constant at its long-run average in our benchmark model. Comparing our full framework to the fixed θ-uncertainty model highlights the specific role played by endogenous uncertainty.

Before comparing the models to the data, we must detrend the empirical time series. We note that the cyclical properties we are interested in, in particular the persistence and depth of the Great

Recession, are sensitive to the specific detrending strategy. In Appendix A.2, we investigate the implications of three filters: the Hodrick and Prescott (1997) (HP) filter, a linear detrending, and a linear detrending allowing for a structural break in the trend. As shown in Figure 12a in that appendix, the HP-filtered data suggests that the Great Recession was a mild economic downturn and that the economy promptly recovered to its long-run trend after the trough. Both conclusions contradict essential features of the raw output data shown in Figure 11a in Appendix A.2.[27] We also find that a purely linear trend exaggerates the severity and persistence of the recession by ignoring low-frequency changes in the trend. Therefore, we choose a linear trend with a structural break estimated by least squares (Hansen, 2000) as our benchmark, and we also report the sensitivity of our results to using a standard linear trend.[28] See Appendix A for a more detailed discussion of the data and the detrending strategies.

We compute standard business cycle moments using the detrended data covering the period between the first quarter of 1978 and the last quarter of 2014. For each of the three models, we generate ten thousand simulations of the same length as the data (148 quarters), and then average each moment across all simulations. Panel A of Table 2 reports the results for the standard deviation of output (Y), consumption (C), employment (L), investment (I), the number of firms (m) and uncertainty about real output growth (U). Panel B reports the correlation between each of these variables and output, and Panel C reports their autocorrelation.

Because of the absence of learning, the RBC model generates more volatility in every variable than our full model. However, their performances are overall similar. Regarding the *volatility* of uncertainty about real output growth, our model is able to explain only a fraction of what is observed in the data, suggesting that other shocks — possibly exogenous uncertainty shocks — may be needed to fully account for the total fluctuations in uncertainty. Note that uncertainty about current-quarter real output growth is 0 in the RBC model under full information. Panels B and C show that the benchmark, the fixed θ-uncertainty model and the RBC model are roughly comparable in terms of the correlations with output and the autocorrelations.

[27] See King and Rebelo (1993) and Cogley and Nason (1995a) for a discussion of various drawbacks of the HP filter.

[28] An additional advantage of the linear trends is that they allow us to cleanly interpret the (unfiltered) prediction of forecasters reported in the SPF data used to calibrate the model. For instance, a 2% growth forecast for this year relative to the previous year can be directly mapped to the model by removing the long-run growth rate in output from the linear filter. It is unclear how the analog exercise would be done if the data is HP filtered and detrended with a bandpass filter. See Appendix A.2 for a detailed discussion of the filters.

	Y	C	L	I	m	U
A. Standard Deviation						
Data	0.039	0.031	0.066	0.132	0.038	0.253
Benchmark	0.035	0.030	0.027	0.072	0.020	0.053
Fixed θ-uncertainty	0.036	0.031	0.027	0.076	0.019	0.027
RBC	0.044	0.036	0.033	0.093	–	–
B. Correlation w.r.t Y						
Data	1	0.588	0.921	0.523	0.646	-0.443
Benchmark	1	0.967	0.884	0.855	0.459	-0.644
Fixed θ-uncertainty	1	0.966	0.888	0.858	0.480	-0.619
RBC	1	0.987	1	0.944	–	–
C. Autocorrelation (1st lag)						
Data	0.981	0.982	0.993	0.962	0.995	0.762
Benchmark	0.956	0.943	0.942	0.918	0.991	0.931
Fixed θ-uncertainty	0.957	0.943	0.941	0.915	0.989	0.869
RBC	0.934	0.949	0.934	0.907	–	–

Notes. All series are computed in log deviation from trend. Each of the 10,000 replications of the simulated series are linearly detrended. The annual series for m is interpolated to quarterly data. Uncertainty U is the standard deviation of current year real output growth expressed in the last quarter of the current year. Because it is only available annually and, by definition, cannot simply be interpolated, the moments we report about uncertainty are computed using observations from the 4th quarter of every year. Its annual autocorrelation is expressed in quarterly terms. C is gross of fixed costs for comparison across models. Employment and output are perfectly correlated in the RBC model as a consequence of the GHH preferences.

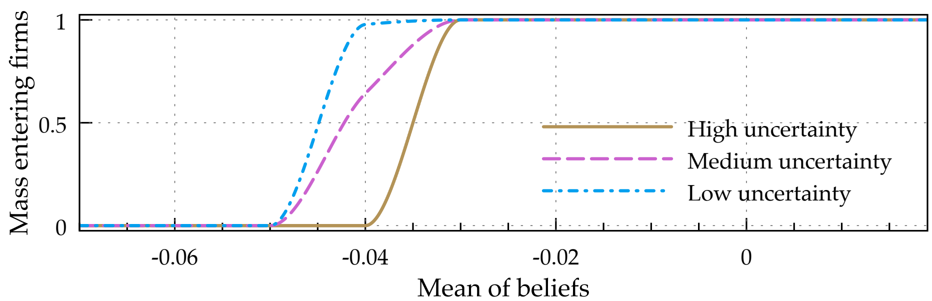
Table 2: Business-Cycle Moments: Data, RBC, and Benchmark Model

The key conclusion from Table 2 is that a standard calibration of our model performs similarly to the RBC model in terms of standard business-cycle moments. Therefore, incorporating endogenous uncertainty in a standard business cycle model does not impair its ability to predict well-known patterns of business cycle data. This result should not be surprising. As we saw in the baseline model, the uncertainty trap mechanism kicks in only for large shocks, which are rare in these simulations. In the next section, we show that the key difference between our model and standard models, and the value added of modeling uncertainty traps, lies in terms of predicting how the economy responds to large shocks.

5.4 Policy Function and Impulse Responses

We now examine the role of endogenous uncertainty in propagating shocks. First, we ask whether the key implication of the uncertainty trap mechanism identified in the baseline theory — generating protracted recessions out of sufficiently negative shocks — is also at work in the full calibrated model. In the next section, we ask whether this feature of the model can help explain the behavior of macro aggregates during the Great Recession.

Firm Entry and Beliefs In Proposition 1 we showed that, under certain conditions, the mass of producing firms increases with the mean of beliefs μ and decreases with uncertainty (increases with γ). We verify that these properties are inherited by the calibrated quantitative model. Figure 8 plots the fraction of entering firms $n(\mu, \gamma, K, m)$ as a function of mean beliefs μ, for three levels of uncertainty. Similarly to Figure 4 in the baseline theory, we find that, for any level of uncertainty, firms are more likely to enter when beliefs are more optimistic and that, for any level of mean beliefs, firms are more likely to enter when uncertainty is lower. Therefore, the entry decision of the quantitative model inherits the key feature of the baseline theory leading to uncertainty traps.

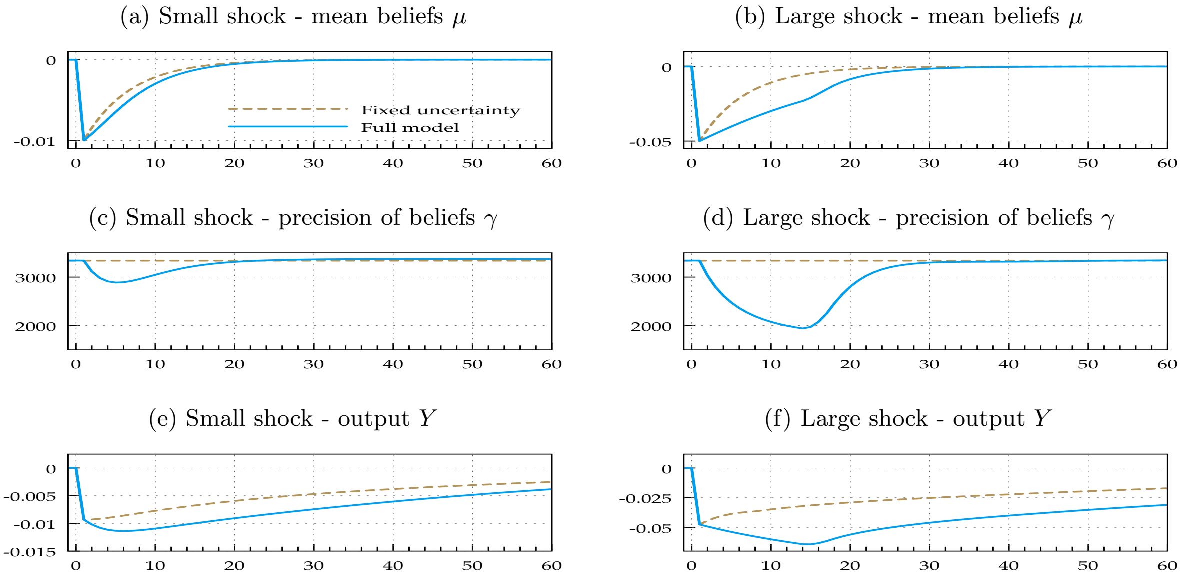


Notes. This figure was generated with K and m at their steady-state values when shocks are set to zero.

Figure 8: Investment decision $n(\mu, \gamma, K, Q)$ for K and m constant at their steady-state level.

Response to Small and Large Shocks A central feature of uncertainty traps is the non-linear response to shocks, shown in Figure 6 in the context of the baseline theory. For shocks leading to large negative drops in the mean of beliefs, endogenous uncertainty leads to recessions which are relatively deeper and last relatively longer than for small shocks. We ask whether the full quantitative model inherits this feature. Figure 9 shows the evolution of the mean of beliefs μ, the precision of beliefs γ, and output Y after a one-period drop in mean beliefs of 1% (left column) and 5% (right column) in both the full model and the fixed θ-uncertainty model.[29] In the fixed-information model, the economy starts to recover immediately after the initial shock regardless of the size of the shock. In contrast, in the full model, output continues to decline after the mean of beliefs has started to recover. Moreover, the duration of this decline is longer the larger is the shock. Reaching the trough in output takes 6 quarters in response to the small shock, and 15 quarters in response to the larger one. As in the baseline model, this non-linearity is driven by the endogenous evolution of uncertainty. The fall in mean beliefs drives down the incentives to enter. As a result, fewer signals are released and the precision of beliefs falls (panels (c) and (d)). After the shock, agents receive signals suggesting an improvement in the fundamental, and beliefs recover (panels (a) and (b)). However, the recovery in output is delayed in the full model by the feedback between high uncertainty (panels (e) and (f)) and slow entry.

[29] Fundamental shocks affect the policy functions only through beliefs. Falls in beliefs may result from shocks to the fundamental θ or to the signals, X and Z.



Notes: The left column shows the response of the economy to a -1% one-period shock to μ. The right column shows the response of the economy to a -5% one-period shock to μ. The solid curves show the evolution of the economy according to the full model, while the dashed curves show the evolution of a control economy in which the flow of public information is fixed at the steady-state level of the full model. Figures (e) and (f) are in log deviation from trend, the other figures are in level.

Figure 9: Mean beliefs, precision of beliefs and output in response to a one-period shock

Table 3 summarizes the properties of the recessions depicted in Figure 9. For both shocks, the table reports the ensuing recession's depth (the magnitude of the peak-to-trough fall in output) and the half-life of the recovery (the number of quarters for the economy to recover half of the peak-to-trough fall in output). The table highlights two key features of uncertainty traps: i) endogenous uncertainty makes recessions deeper and longer for shocks of any magnitude (recessions' depth and duration are larger in the full than in the fixed θ-uncertainty model), and ii) the differential effect of endogenous uncertainty is relatively larger for large than for small shocks. The recession is 22% deeper and 40% longer in the full model than in the fixed θ-uncertainty model under a negative 1% shock to beliefs, but 35% deeper and 66% longer under a negative 5% shock. Therefore, in the full calibrated model, endogenous uncertainty leads to amplification and persistence of shocks driving down beliefs.

	Small shock (-1%)		Large shock (-5%)	
	Depth	50% recovery	Depth	50% recovery
Full model	-1.1%	42	-6.2%	58
Fixed θ-uncertainty	-0.9%	30	-4.6%	35

Notes. The depth of the recession corresponds to the lowest value of output reached since the official beginning of the recession. The "50% recovery" column is the number of quarters before the economy recovers 50% of the peak-to-trough drop in output.

Table 3: The impact of shock sizes on the depth and duration of recessions across models

5.5 Endogenous Uncertainty in U.S. Recessions

Our previous discussion established that, within our calibrated model, endogenous uncertainty amplifies and lengthens the decline in economic activity relatively more for large than for small shocks. We now evaluate the role of endogenous uncertainty in explaining the U.S. experience during past recessions. Specifically, we ask whether it can help explain the observed depth and persistence of a recession when there is a large enough shock and if, indeed, the mechanism only "kicks in" for sufficiently large shocks.

To implement this exercise we must first take a stand on the exogenous shocks that hit the U.S. economy in each recession. Within our model, the only exogenous shocks are the innovations to the signals $\{X_t, Z_t\}$ and to the fundamentals $\{\theta_t\}$. The shocks to fundamental $\{\theta_t\}$ are taken directly from the data. To discipline the evolution of $\{X_t, Z_t\}$, we choose their innovations such that our model and the fixed θ-uncertainty model exactly reproduce, as equilibrium outcomes, the forecast of current year's real output from the SPF during the first part of each recession.[30] We then compute the response of the three models we have introduced – our full quantitative model, the fixed θ-uncertainty model, and the RBC model – and contrast them with the data. Importantly, among the variables that we simulate and contrast against the data we include uncertainty about output growth from the SPF and the number of firms, both of which are directly related to the forces in the model.

We implement this exercise in the context of the largest and second-largest recessions in terms of decline in output within our sample, the Great Recession and the 1981-1982 recession. Since we expect uncertainty traps to be more important after large shocks, the Great Recession — the most important downturn since the Great Depression — is a historical episode where we would expect endogenous uncertainty to play a role. However, we also verify that endogenous uncertainty does not transform every downturn into a deeply protracted recession. For that, we also consider the impact of the mechanism during the 1981-1982 recession, which was characterized by a large but brief decline in productivity and was followed by a relatively rapid recovery in output. We expect the impact of endogenous uncertainty to be relatively weaker during this episode.

5.5.1 Great Recession

We start by considering the Great Recession. According to the NBER, the recession took place over the period 2007:Q4-2009:Q2. We first feed the model with the observed TFP shocks over the corresponding period.[31] We then feed each model with a series of signals that allows them to

[30] Note that we do not need to specify exactly the combination of X_t and Z_t shocks that is needed to match output forecasts. Only the time series of μ_t is relevant for our purpose. In the case of the RBC model we only introduce shocks to θ_t, as there is no degree of freedom to match the forecast about real output.

[31] We pick 2007:Q2 as our start date to be consistent with the dates at which we match output forecasts. However, we stop feeding the series in 2008:Q4 because the TFP series from Fernald (2014) displays a large increase at the beginning of 2009, which leads to a strong counterfactual expansion in the RBC economy (the other models fare better). For robustness, we fit the full TFP series in Figure 13 of Appendix C and experiment with other TFP series in Figures 15 and 16. In all cases, the full model generates a deeper and longer recession than the fixed θ-uncertainty and RBC models.

replicate output forecasts from the SPF over that period. Specifically, the SPF surveys forecasters every quarter about their point expectations for the level of real output for the current year. Since the trough in output forecast from the SPF was reached in 2009:Q2, we focus on second quarter forecasts and match expectations about current year output in 2007:Q2, 2008:Q2, and 2009:Q2.[32]

Figure 10 shows the evolution of TFP, real output forecasts, output, investment, employment, consumption, the number of firms, and uncertainty about real GDP growth. The solid lines in the left column represent the data. The right column shows the predictions of the three models: the full quantitative model (solid blue curves), the fixed θ-uncertainty model (dashed brown curves), and the RBC model (dot-dash pink curves). The diamonds in panel (b) and (d) of Figure 10 indicate the data that the models are calibrated to match. All shocks and innovations are set to zero after 2009:Q2, and the economy is left to recover.

The central finding of this exercise, is that for all macro aggregates that we consider (output, investment, consumption, employment, number of firms and uncertainty about GDP growth), our model generates dynamics which are closer to the data than the fixed θ-uncertainty model and the RBC model. Panel (f) of Figure 10 shows that the decline in output is clearly deeper and more protracted in our model than in the alternative models.

Table 4 summarizes the key properties of the recession in the data and as predicted by the three models. The first column shows the depth of the recession, measured as the percentage drop in output from the beginning of the recession to its lowest value. The full model is able to explain almost all of the depth, while the fixed θ-uncertainty model and the RBC model account for 76% and 33% of the fall, respectively. The second column of the table shows the earliest date at which output has recovered 20% of the depth.[33] In the full model, output recovers 20% of the lost ground in 2015:Q1, close to corresponding date in the data (2014:Q3). In contrast, in the fixed θ-uncertainty and in the RBC models this event happens in 2012:Q4 and 2009:Q3, respectively. These results imply that, within our model, endogenous uncertainty adds 1.8 (5.2) percentage points in terms of depth and about two (five) years until the 20% recovery relative to the fixed θ-uncertainty (RBC) model. Finally, the third column shows the cumulative loss in output, measured as the sum of the percentage deviations from trend from 2007:Q4 to 2015:Q1, the end of our sample. This measure combines the depth and the persistence of the recession into a single statistic. If we normalize this measure by the output produced in quarter 2007:Q4 at the onset of the recession, the U.S. economy lost a cumulated 1.84 quarters of output since that date. Our full model explains 93% of this loss, while the fixed θ-uncertainty and RBC models can account for 73% and 30% of the loss,

[32] Forecasts in the first, third or fourth quarters in the SPF are about the current year and thus correspond to forecasts at different horizons. We focus on a single horizon with second-quarter forecasts to avoid using four different definitions of forecast in our model. Since we do not use forecast data in Q1, Q3 and Q4, we linearly interpolate the series of μ_t at these points. Note finally that the empirical output forecast series stops in 2009:Q2 because of a change in variable definition.

[33] We cannot use the standard half-life measure of persistence since output has not recovered half of the way from its trough in the data. We pick the 20% threshold as it roughly corresponds to the level attained by output in the data at the end of the sample. The full model also generates a longer recession measured in terms of the number of quarters from peak to trough relative to alternative models. The exact date for the trough is however hard to identify in the data because of high-frequency noise.

respectively. According to these various measures, our full model is better able to account for the overall severity of the recession.

	Depth	20% recovery	Cumulative loss
Data	-7.9%	2014:Q3	-184%
Full model	-7.8%	2015:Q1	-171%
Fixed θ-uncertainty	-6.0%	2012:Q4	-134%
RBC	-2.6%	2009:Q3	-56%

Notes. We compute the depth, persistence, and cumulative loss of the recession from the official beginning of the recession in 2007:Q4 as determined by the NBER. The depth of the recession corresponds to the lowest value of output reached since the official beginning of the recession. The "20% recovery" column is the earliest date at which output recovers 20% of the depth. Cumulative loss is the sum of the percentage deviation of output from trend from 2007:Q4 to 2015:Q1.

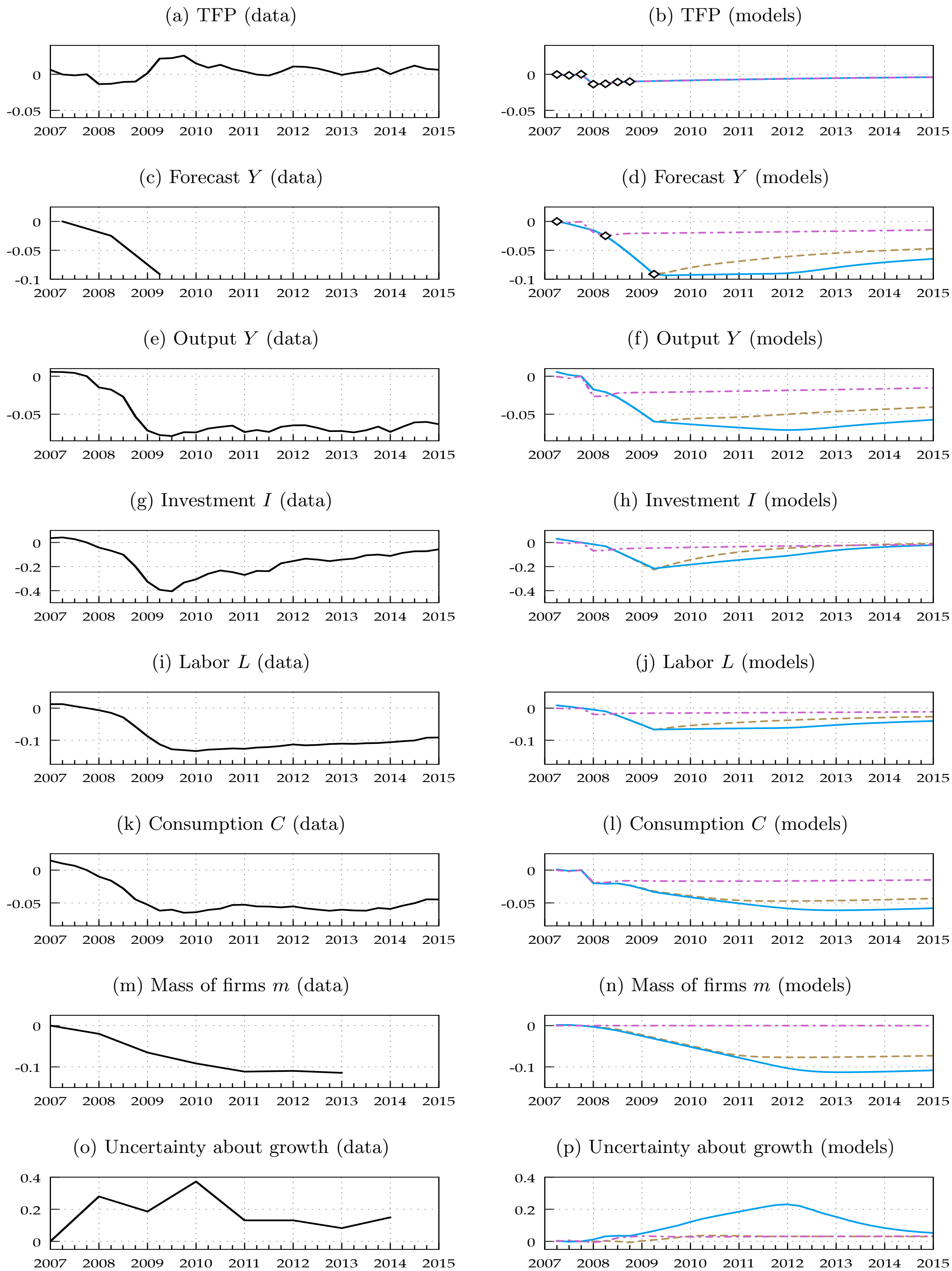
Table 4: Depth, persistence, and cumulative output loss of the Great Recession

Besides output, endogenous uncertainty can also help explain the behavior of other aggregates. As shown in panels (g) and (h), even though our model does not generate as pronounced a decline in investment, it does generate more persistence than the alternative models. Panels (m) and (n) of Figure 10 show that the number of firms also falls. By construction, the RBC model does not have predictions in this regard. Our full quantitative model does well in predicting the persistence and the amplitude of the fall in the number of active firms. Since entry and the number of firms govern the dynamics of uncertainty within the full model, the good performance of the model along that dimension is reassuring. Our model generates as large a fall in consumption as it is observed in the data (panels (k) and (l)). Relative to the alternative models, it also predicts a larger fall and a more protracted recovery in employment (panels (i) and (j)), better reflecting the empirical patterns. Even though for most variables the fall in economic aggregates happens faster in the data, our model is closer to the empirical series than the alternative models for all the variables reported in the figure.

We also highlight that our model is relatively successful at explaining the shape of the uncertainty response. As observed in panels (o) and (p) of Figure 10, the uncertainty in output growth follows an inverted-U shape relationship that our model reproduces, although in our model the peak in uncertainty is delayed relative to the data. In contrast, the alternative models generate little variation in uncertainty in GDP forecasts.

Sensitivity

We demonstrate the robustness of our findings by replicating the Great Recession exercise under alternative assumptions: i) fitting the whole TFP series instead of only its beginning; ii) using a simple linear detrending instead of allowing for a structural break; iii) using TFP data from the Penn World Table and the Bureau of Labor Statistics instead of the Fernald time series; and *iv)* using



Notes. The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

Figure 10: Great Recession - Data and models

standard CRRA preferences instead of the GHH preferences assumed in the main exercise. The full simulations from these robustness exercises are included in Appendix C. Table 5 summarizes these simulations by showing the cumulative loss in output generated by each model in each exercise, and comparing it to the data. In all cases, we find that our conclusions are robust. The full model with endogenous uncertainty gets closer to the data by generating more severe recessions than the alternative models. Table 7 in the Appendix shows that the full model generates deeper and longer lasting recessions in all the robustness exercises.

	Cumulative loss in output			
	Data	Full model	Fixed θ-uncertainty	RBC
Benchmark	-184%	-166%	-129%	-56%
Full TFP fit	-184%	-86%	-38%	+38%
Linear detrending	-310%	-173%	-139%	-72%
PWT TFP	-184%	-194%	-175%	-131%
BLS TFP	-184%	-171%	-142%	-85%
CRRA Preferences	-184%	-80%	-65%	-31%

Notes. Cumulative loss is the sum of the percentage deviation of output from trend from 2007:Q4 to 2015:Q1.

Table 5: Robustness: Cumulative loss during the Great Recession

5.5.2 1981-1982 Recession

We now consider the impact of endogenous uncertainty in the less severe recession of 1981Q3-1982Q4. The purpose of this section is to show that, while endogenous uncertainty generates amplification and persistence in response to very large shocks, it does not necessarily increase the depth and persistence of every negative shock. As such, the model does not have counterfactual implications for the magnitude and persistence of milder recessions.

Table 6 replicates Table 4 for the 1981-1982 recession.[34] We see that the full and the fixed θ-uncertainty models generate shallower recessions than the data, while output drops more than the data in the RBC model. In terms of persistence, the three models performs similarly, suggesting that the economy should have recovered 20% of the loss in output at the beginning of 1984 — about a year after that threshold was reached in the data. In terms of cumulative loss, all models predict recessions that are more severe than the data, although the difference between model and data is larger for the RBC than for the full and fixed θ-uncertainty models. Overall, we find that for the 1981-1982 recession, characterized by a less severe combination of TFP and beliefs shocks, endogenous uncertainty does not generate additional persistence. This confirms that the full model generally behaves as a standard RBC model, except during large unusual downturns, such as the 2007-2009 recession, when the uncertainty trap mechanism kicks in.

[34] Since the output forecast series is fully available over the period 1981-1985, we fit the whole time series of both beliefs and TFP. Appendix D shows the model predicted series.

	Depth	20% recovery	Cumulative loss
Data	-5.3%	1983:Q2	-13%
Full model	-4.9%	1984:Q1	-32%
Fixed θ-uncertainty	-4.7%	1984:Q1	-32%
RBC	-6.4%	1984:Q2	-44%

Notes. The table is constructed similarly to Table 4. The official beginning of the recession is in 1981:Q3, and the cumulative loss is computed between that quarter and 1985:Q1.

Table 6: Depth, persistence and cumulative output loss of the 1981-1982 recession

5.6 Further Applications

In Appendix E we provide additional exercises that highlight various features of the model. First, we show that, like other studies of the effects of uncertainty on the economy (Bloom, 2009; Bloom et al., 2012), exogenous uncertainty shocks generate recessions in our framework. In our model, these exogenous shocks are propagated endogenously through the endogenous uncertainty channel.

Second, we consider the problem of a social planner in Appendix E.2. To do so, we compute the efficient allocation and consider how it responds to shocks relative to the competitive equilibrium that we have studied so far. We find that the planner prevents the number of entering firms from falling too much during the recession in order to prevent uncertainty from rising too much. As a result, the recession is shorter and shallower in the planner's allocation, suggesting a potential role for government policies in mitigating the impact of recessions.

Third, we show in Appendix E.3 that our benchmark model improves on RBC in terms of propagation of shocks. As noted by Cogley and Nason (1995b), the RBC model features weak internal persistence so that the properties of GDP growth mimic those of the exogenous TFP growth process. In our full model, however, the persistence in output growth is larger than in TFP growth, and more so at lower lags. Hence, while our model does not replicate the full amount of autocorrelation observed in the data, it does reproduce the *qualitative* feature of the data that the autocorrelation of GDP growth is larger than that of TFP growth at lower lags. We interpret this finding as evidence that endogenous uncertainty leads to a more persistent series for output growth relative to TFP growth, in line with the baseline theory.

6 Conclusion

We develop a theory of endogenous uncertainty and business cycles that combines two forces: higher uncertainty about economic fundamentals deters investment, and uncertainty evolves endogenously because agents learn from the actions of others. The interaction between investment and uncertainty leads to *uncertainty traps*: episodes in which high uncertainty leads firms to delay investment, further raising uncertainty. In the unique equilibrium of the model, the economy fluctuates between a high-activity/low-uncertainty regime and a low-activity/high-uncertainty regime and is subject to strong non-linear dynamics in which large shocks can have near permanent effects.

To quantify the importance of uncertainty traps, we embed the mechanism in a business cycle model. We calibrate the model to the United States economy and find that endogenous uncertainty increases the depth and duration of recessions, particularly after large shocks. We also find that our full model is able to better explain the behavior of standard macro aggregates during the Great Recession than alternative models.

We believe that the novel channel proposed in this paper is important for several reasons. First, the emphasis on subjective uncertainty — and beliefs about fundamentals in particular — implies that not only exogenous volatility shocks, but also other sources of uncertainty, matter for the economy. Thus, we view recent empirical work using survey data on forecasts or consumer and business expectations as an important step towards a more complete understanding of the role of uncertainty in business cycles. Second, we believe that our framework may be useful as a theoretical benchmark for empirical and quantitative studies seeking to estimate the direct and feedback effects of uncertainty on economic activity. Despite the multiplicity of regimes and strong non-linearities, the model features a single competitive equilibrium, which makes it amenable to applied work. Third, we have shown that allowing uncertainty to fluctuate endogenously leads to a significant propagation and amplification mechanism. The type of non-linearities and the multiplicity in regimes that we obtain may be of broader interest for business cycle modeling in general and could also shed light on some particularly large historical downturns.

For the sake of clarity, we have exposited the mechanism in a purposely simple framework, but a number of generalizations may be worth investigating. In particular, it would be interesting to understand how uncertainty traps interact with frictions that could magnify their impact, such as financial frictions, demand externalities, or belief heterogeneity. We leave these questions to future research.

References

Amador, M. and P.-O. Weill (2010): "Learning from Prices: Public Communication and Welfare," *Journal of Political Economy*, 118.

Angeletos, G.-M., C. Hellwig, and A. Pavan (2007): "Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks," *Econometrica*, 75, 711–756.

Arellano, C., Y. Bai, and P. Kehoe (2012): "Financial Markets and Fluctuations in Uncertainty," Working paper, Federal Reserve Bank of Minneapolis and NBER; Arizona State University; University of Minnesota and Federal Reserve Bank of Minneapolis.

Bachmann, R. and C. Bayer (2013): "Wait-and-See Business Cycles?" *Journal of Monetary Economics*, 60, 704–719.

——— (2014): "Investment Dispersion and the Business Cycle," *The American Economic Review*, 104, pp. 1392–1416.

Bachmann, R., S. Elstner, and E. R. Sims (2013): "Uncertainty and Economic Activity: Evidence from Business Survey Data," *American Economic Journal: Macroeconomics*, 5, 217–49.

Bachmann, R. and G. Moscarini (2011): "Business cycles and endogenous uncertainty," *manuscript, Yale University, July.*

Banerjee, A. V. (1992): "A simple model of herd behavior," *The Quarterly Journal of Economics*, 107, 797–817.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 823–866.

Barro, R. J. and J. F. Ursúa (2012): "Rare Macroeconomic Disasters," *Annual Review of Economics*, 4, 83–109.

Basu, S. and J. G. Fernald (1997): "Returns to scale in US production: Estimates and implications," *Journal of political economy*, 105, 249–283.

Beaudry, P. and F. Portier (2006): "Stock Prices, News, and Economic Fluctuations," *The American Economic Review*, 96, 1293–1307.

Benhabib, J., P. Wang, and Y. Wen (2015): "Sentiments and Aggregate Demand Fluctuations," *Econometrica*, 83, 549–585.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 992–1026.

——— (1998): "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *The Journal of Economic Perspectives*, 12, pp. 151–170.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. Terry (2012): "Really Uncertain Business Cycles," Nber working paper no.18245, NBER.

Canova, F. (1998): "Detrending and business cycle facts," *Journal of Monetary Economics*, 41, 475 – 512.

Caplin, A. and J. Leahy (1993): "Sectoral shocks, learning, and aggregate fluctuations," *The Review of Economic Studies*, 60, 777–794.

Carlsson, M. (2007): "Investment and Uncertainty: A Theory-based Empirical Approach," *Oxford Bulletin of Economics and Statistics*, 69, 603–617.

Chamley, C. (1999): "Coordinating regime switches," *The Quarterly Journal of Economics*, 114, 869–905.

——— (2004): "Delays and equilibria with large and small information in social learning," *European Economic Review*, 48, 477 – 501.

Chamley, C. and D. Gale (1994): "Information revelation and strategic delay in a model of investment," *Econometrica: Journal of the Econometric Society*, 1065–1085.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011): "Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins," *The American Economic Review*, 101, 471–475.

Cogley, T. and J. M. Nason (1995a): "Effects of the Hodrick-Prescott filter on trend and difference stationary time series Implications for business cycle research," *Journal of Economic Dynamics and control*, 19, 253–278.

——— (1995b): "Output Dynamics in Real-Business-Cycle Models," *The American Economic Review*, 85, pp. 492–511.

Decker, R. A. and P. N. D'Erasmo (2016): "Market exposure and endogenous firm volatility over the business cycle," *American Economic Journal: Macroeconomics*, 8, 148–198.

Dixit, A. K. and R. S. Pindyck (1994): *Investment under uncertainty*, Princeton university press.

Dontchev, A. L. and R. T. Rockafellar (2009): "Implicit functions and solution mappings," *Springer Monographs in Mathematics. Springer*, 194.

Farmer, R. (2013): "Animal Spirits, Financial Crises and Persistent Unemployment," *Economic Journal*, 123, 317–340.

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015): "The Next Generation of the Penn World Table," *American Economic Review*, 105, 3150–3182.

Fernald, J. G. (2014): "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," *Manuscript, Federal Reserve Bank of San Francisco.*

Foster, A. D. and M. R. Rosenzweig (1995): "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103, pp. 1176–1209.

Gabaix, X. (2012): "Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance*," *The Quarterly journal of economics*, 127, 645–700.

Gilchrist, S., J. W. Sim, and E. Zakrajšek (2014): "Uncertainty, financial frictions, and investment dynamics," Tech. rep., National Bureau of Economic Research.

Gourio, F. (2012): "Disaster Risk and Business Cycles," *American Economic Review*, 102, 2734–66.

Guiso, L. and F. Schivardi (2007): "Spillovers in Industrial Districts," *The Economic Journal*, 117, pp. 68–93.

Hansen, B. E. (2000): "Testing for structural change in conditional models," *Journal of Econometrics*, 97, 93–115.

Hodrick, R. J. and E. C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29, 1–16.

Jaimovich, N. and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle?" *American Economic Review*, 99, 1097–1118.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," *American Econonmic Review*, 105, 1177–1216.

Kaplan, G. and G. Menzio (2013): "Shopping Externalities and Self-Fulfilling Unemployment Fluctuations," .

Kehrig, M. (2011): "The cyclicality of productivity dispersion," *US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15.*

Kim, T. (1993): "Differentiability of the Value Function," .

King, R. G. and S. T. Rebelo (1993): "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control*, 17, 207 – 231.

KOZLOWSKI, J., L. VELDKAMP, AND V. VENKATESWARAN (2015): "The Tail that Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation," Working paper.

LANG, W. W. AND L. I. NAKAMURA (1990): "The dynamics of credit markets in a model with learning," *Journal of Monetary Economics*, 26, 305 – 318.

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

LUCAS JR, R. E. (1978): "On the size distribution of business firms," *The Bell Journal of Economics*, 508–523.

MORRIS, S. AND H. S. SHIN (1998): "Unique equilibrium in a model of self-fulfilling currency attacks," *American Economic Review*, 587–597.

——— (1999): "A theory of the onset of currency attacks," *The Asian Financial Crisis*, 230.

ORDOÑEZ, G. L. (2009): "Larger crises, slower recoveries: the asymmetric effects of financial frictions," Tech. rep., Federal Reserve Bank of Minneapolis.

ORLIK, A. AND L. VELDKAMP (2013): "Understanding Uncertainty Shocks and the Role of Black Swans," working paper.

RAVN, M. O. AND H. UHLIG (2002): "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *The Review of Economics and Statistics*, 84, 371–376.

ROB, R. (1991): "Learning and capacity expansion under demand uncertainty," *The Review of Economic Studies*, 58, 655–675.

SCHAAL, E. (2015): "Uncertainty and Unemployment," working paper, New York University.

SCHAAL, E. AND M. TASCHEREAU-DUMOUCHEL (2015): "Coordinating Business Cycles," working paper.

STOKEY, N. L. (2008): *The Economics of Inaction: Stochastic Control models with fixed costs*, Princeton University Press.

STOKEY, N. L., E. C. PRESCOTT, AND R. E. LUCAS (1989): *Recursive methods in economic dynamics / Nancy L. Stokey and Robert E. Lucas, Jr., with Edward C. Prescott*, Harvard University Press Cambridge, Mass.

STRAUB, L. AND R. ULBRICHT (2015): "Endogenous Uncertainty and Credit Crunches," *manuscript*.

VAN NIEUWERBURGH, S. AND L. VELDKAMP (2006): "Learning asymmetries in real business cycles," *Journal of Monetary Economics*, 53, 753–772.

Vavra, J. (2014): "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation*." *Quarterly Journal of Economics*, 129.

Veldkamp, L. L. (2005): "Slow boom, sudden crash," *Journal of Economic Theory*, 124, 230–257.

Zeira, J. (1994): "Informational cycles," *The Review of Economic Studies*, 61, 31–44.

A Data and detrending

This appendix details the sources of data used throughout the paper and the way in which the data is detrended.

A.1 The sources of data

This section lists the data sources that we use for the measurement of uncertainty and for the quantitative exercises. As one the main time series we use, the number of firms, begins in 1978, we limit our analysis to the post-1978 period.

- Our benchmark series for Total Factor Productivity in the US is the quarterly Business Sector TFP, adjusted for capacity utilization, from Fernald (2014) over the period 1978:Q1-2015:Q1 and seasonally adjusted.

- From the Bureau of Economic Analysis, we use the quarterly Real Gross Domestic Product (Fred: GDPC1), Real Gross Private Domestic Investment (Fred: GPDIC1) and the Real Personal Consumption Expenditure (Fred: PCECC96) over the period 1978:Q1-2015:Q3 and seasonally adjusted.

- Our measure of total labor is the quarterly Nonfarm Business Sector: Hours of All Persons (Fred: HOANBS) measure from the Bureau of Labor statistics over the period 1978:Q1-2015:Q1 and seasonally adjusted.

- For data about firm dynamics, we use the Business Dynamics Statistics from the Census Bureau. This data is available at an annual frequency from 1978 to 2013.

- We use two series from the Survey of Professional Forecasters (SPF). First, as a proxy for uncertainty, we use the mean probability forecasts (PRGDP) which provides an average across forecasters for the probability that the current annual-average over past annual-average percent change in real GDP falls in a particular bin. We compute the standard deviation of this probability distribution. The time series covers the period 1992:Q1-2015:Q1. Second, to calibrate the beliefs during the Great Recession, we use the mean forecasts for the level of real GDP.

- We use the uncertainty series H1, H3 and H12 constructed by Jurado et al. (2015) over the period 1960:Q3-2014:Q4.

- We use the VXO time series from the Chicago Board Options Exchange over the period 1986:Q1-2014:Q2 averaged over monthly periods.

- We use the Michigan Survey of Consumers over the period 1960:Q1-2015:Q1 and construct series with the number of respondents answering "Uncertain Future" as the main reason why it is a bad time to purchase big household goods.

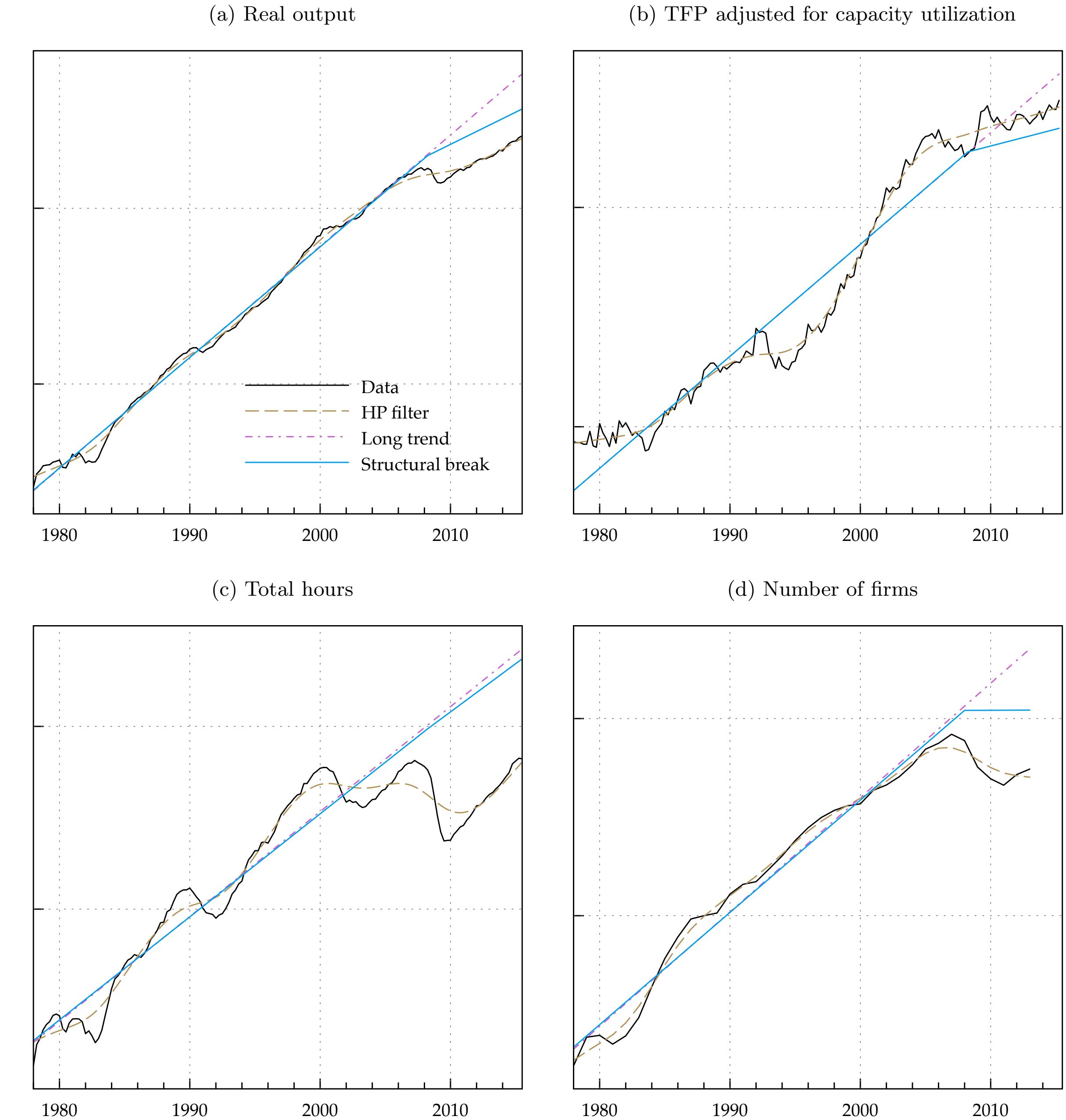
- To verify the robustness of our empirical exercise about the 2007-2009 US recession we also use two alternative measures of total factor productivity. The first one is the Penn World Table Total Factor Productivity at Constant National Prices for the United States (Fred: RTFPNAUSA632NRUG). This data is currently only available at an annual frequency and until 2011. The second productivity times series we use is the Bureau of Labor Statistics Net Multifactor Productivity for the Private Business Sector. This data is currently only available at an annual frequency from 1987 to 2014.

A.2 Discussion of the Detrending Strategy

We consider how alternative ways of detrending the data affect the observed persistence of the Great Recession and discuss our choice for the filter used in Sections 5.3 and 5.5.

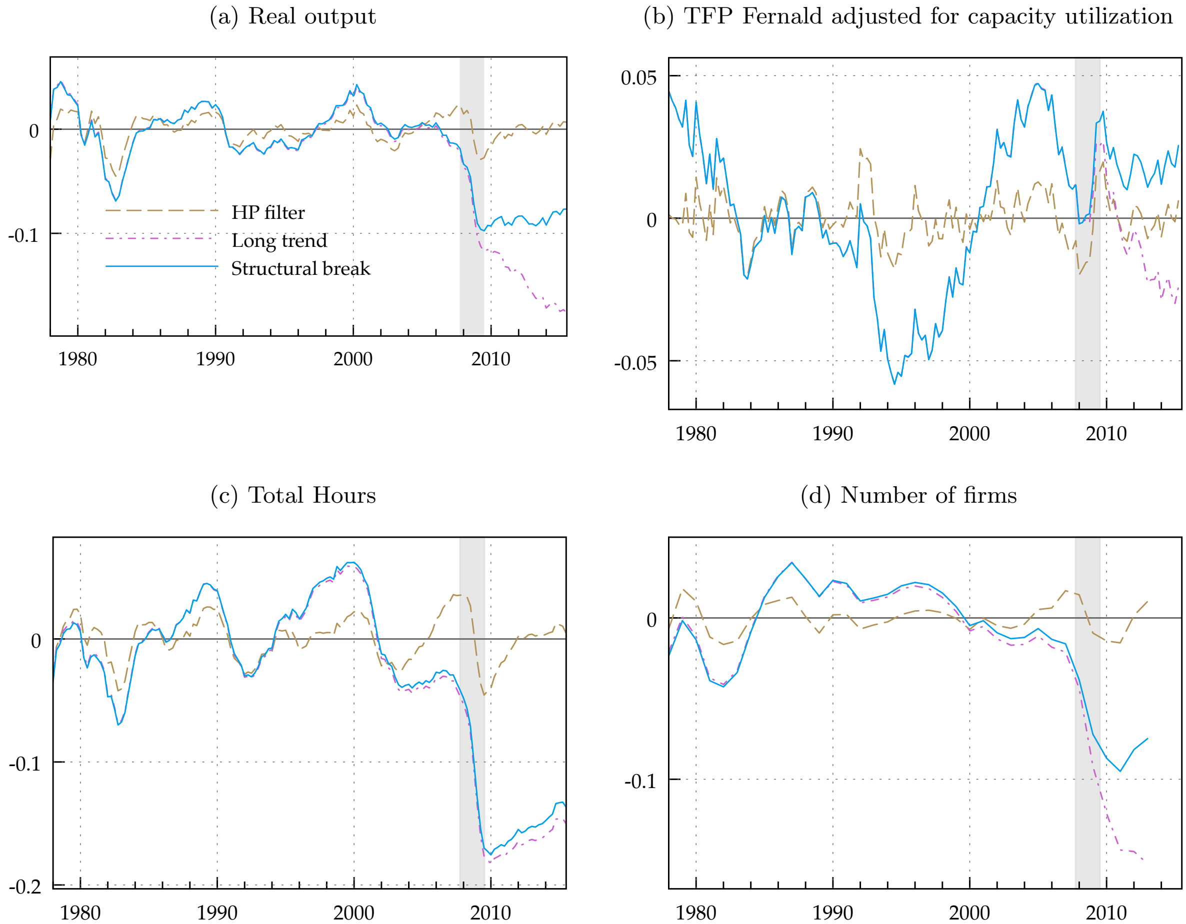
Figure 11 shows four of the main data series used in Section 5: output, TFP, labor and the number of firms. For each series, we show the trends generated by three different filters: i) the brown dashed line is the trend generated by a standard Hodrick and Prescott (1997) filter;[35] ii) the dot-dashed pink line is the best-fitting linear trend between 1978:Q1 and the last quarter before the beginning of the Great Recession in 2007:Q3; and iii) the solid blue line is the best linear fit generated by a least-square estimation with a structural break in the coefficients (see details below). Figure 12 shows the detrended data corresponding to each trend.

[35] We follow Ravn and Uhlig (2002) and use a filter with smoothing parameter of 1600 for quarterly data and 6.25 for annual data. A bandpass filter provides results similar to the HP filter (see Canova (1998)).



Notes: The black solid line represents the data. The solid blue line is the trend from the structural break estimation. The dashed blue line represents the estimated fit of the structural break. The brown dashed line represents the trend from the Hodrick and Prescott (1997) filter with parameter 1600 for quarterly data and 100 for annual data.

Figure 11: Raw data and the trends from various filtering



Notes: The pink dashed line represents the data detrended with the linear trend from 1978:Q1 to 2007:Q3. The solid blue line represents the data detrended with the structural break estimation. The brown dashed line represents the data detrended with the Hodrick and Prescott (1997) filter with parameter 1600 for quarterly data and 100 for annual data. The shaded area corresponds to the 2007Q4 - 2009:Q2 NBER recession

Figure 12: The data detrended by various filters

A comparison between raw output data in Figure 11a and the HP-filtered output data in Figure 12a suggests that the HP-filtered data does not properly reflect either the depth or the persistence of the Great Recession. The HP-filtered data implies that the Great Recession was a mild economic downturn and that the economy promptly recovered to its long-run state. These features of the HP-filtered data clearly miss the important characteristics of the Great Recession that make it unique by historical standards, as the raw data shows in Figure 12a.[36] Therefore, we consider that the alternative linear filters better reflect the depth and persistence of the recession observed in the raw data.

[36] See King and Rebelo (1993) and Cogley and Nason (1995a) for a discussion of various drawbacks of the HP filter.

The raw data also suggests the possibility of a change in the trend at the time of the Great Recession. This observation is confirmed by the fact that the economy keeps moving away from its historical linear trend over the period 2009-2015, as seen in Panel (a). We therefore consider a third linear filter that authorizes a break in the trend. The estimation (Hansen, 2000) finds that a structural break in output occurred during the second quarter of 2008. We use the same break point to detrend the other time series.[37] This filtering strategy delivers a middle ground between the very fast recovery suggested by the HP filter and the very protracted and ongoing decline implied by the linear trend. We thus use the linear detrending with a structural break as our benchmark, and we replicate the Great-Recession exercise using the simple linear trend in one of the sensitivity checks of Section C.

A.3 Filters used in various part of the paper

- In accordance with our calibration strategy which targets moments over 1978:Q1 to 2007:Q3, we use the Fernald TFP data adjusted for capacity utilization and linearly detrended over the same period.

- For the benchmark Great Recession exercise of Section 5.5 we use the data detrended with the structural break approach described in the previous section. We compute the linear trend of the data before and after the 2008:Q2 structural break estimated from the output time series. Using balanced-growth path arguments, we detrend consumption and investment using the same trend as output.[38] Since the data about the number of firms is at an annual frequency and that very few data points are after the break date, we detrend the number of firms using the linear trend from 1978:Q1 to 2007:Q3.[39]

- In the robustness exercises of section C in which we use different TFP series, we detrend all series as in the benchmark exercise except the PWT and BLS data. These data are annual and have very few data points after the structural break such any trend estimation would be meaningless. We instead detrend these data using the 1978:Q1-2007:Q3 linear trend.

[37] As a robustness, we have also estimated structural break points independently on each variable. In most cases, the estimation suggests that the break occurred around the time of the recession. An exception is the time series for hours, which implies a break around 2000. We do not use the structural break detrending for the number of firms, the Penn World Table TFP and the Bureau of Labor Statistics TFP. See the next section for details.

[38] By the resource constraint, the long-run growth rate of consumption and investment should be the same. We have also done the whole Great Recession exercise with consumption and investment data detrended independently. The differences with the benchmark are small. In both case, the mechanism generates substantial amplification and persistence over the RBC and fixed θ-uncertainty models.

[39] We have also replicated the Great Recession exercise using the number of firms data detrended with the structural break procedure. The only difference is that m recuperates slightly in this case, as seen in figure 12d. The mechanism still generates substantial amplification and persistence over the RBC and fixed θ-uncertainty models.

B Numerical Algorithm

This appendix describes the numerical algorithm used to compute the equilibrium in the quantitative model. The algorithm combines two loops. The outer loop iterates over the laws of motion for beliefs and the aggregate mass of entering firms. The inner loop iterates over the capital accumulation decision of the firms. We index outer loop iterations by k and inner loop iterations by l. The steps are as follows:

1. Initialize $k = 0$.

2. Guess a law of motion for entry $n_{(k)}\left(\mu, \gamma, K, m\right)$

3. Beginning of the outer loop

   (a) From $n_{(k)}$, compute the laws of motion of m and of the beliefs μ' and γ'.

   (b) Beginning of the inner loop. Set $l = 0$.

      i. Guess aggregate consumption $C_{(l)}\left(\mu, \gamma, K, m, \theta, f\right)$

      ii. Compute the right-hand-side of the firms' first-order condition for capital accumulation as a function of K'[40]

      $$\text{RHS}\left(K'\right) = \beta E\left[U_c\left(C'_{(l)}, L'\right)\left(\alpha\omega A\left(1+\theta'\right) m^{1-\omega} K'^{\alpha\omega-1} L'^{(1-\alpha)\omega} + 1 - \delta_K\right) \mid \mu, \gamma\right]$$

      iii. Using a nonlinear solver, find the value K' that solves the firms' first-order condition for capital accumulation

      $$E\left[U_c\left(A\left(1+\theta\right) K^{\alpha\omega} L^{(1-\alpha)\omega} + \left(1-\delta_K\right) K - mf - K', L\right) \mid \mu, \gamma\right] = \text{RHS}\left(K'\right)$$

      iv. Using the resource constraint, compute the corresponding current consumption $C_{(l+1)}$, move to the next inner iteration $l \to l+1$ and return to step *i)* with the new guess for consumption. Repeat until C converges.

   (c) Compute the value of waiting and the value of entering for the firms. Find the equilibrium entry decision $n_{(k+1)}\left(\mu, \gamma, K, m\right)$. Move to the next outer iteration $k \to k+1$ and return to step *a)* with the new guess on n.

   (d) Repeat until n converges.

C Sensitivity Analysis on the Great Recession Exercise

This appendix evaluates the robustness of the Great Recession exercise of Section 5 by relaxing or modifying various assumptions. In Figure 13, we repeat the exercise by matching the full time

[40]GHH preferences give us an analytical solution from labor supply as a function of the state space. Also, since all firms make the same decision we already use $K = mk$.

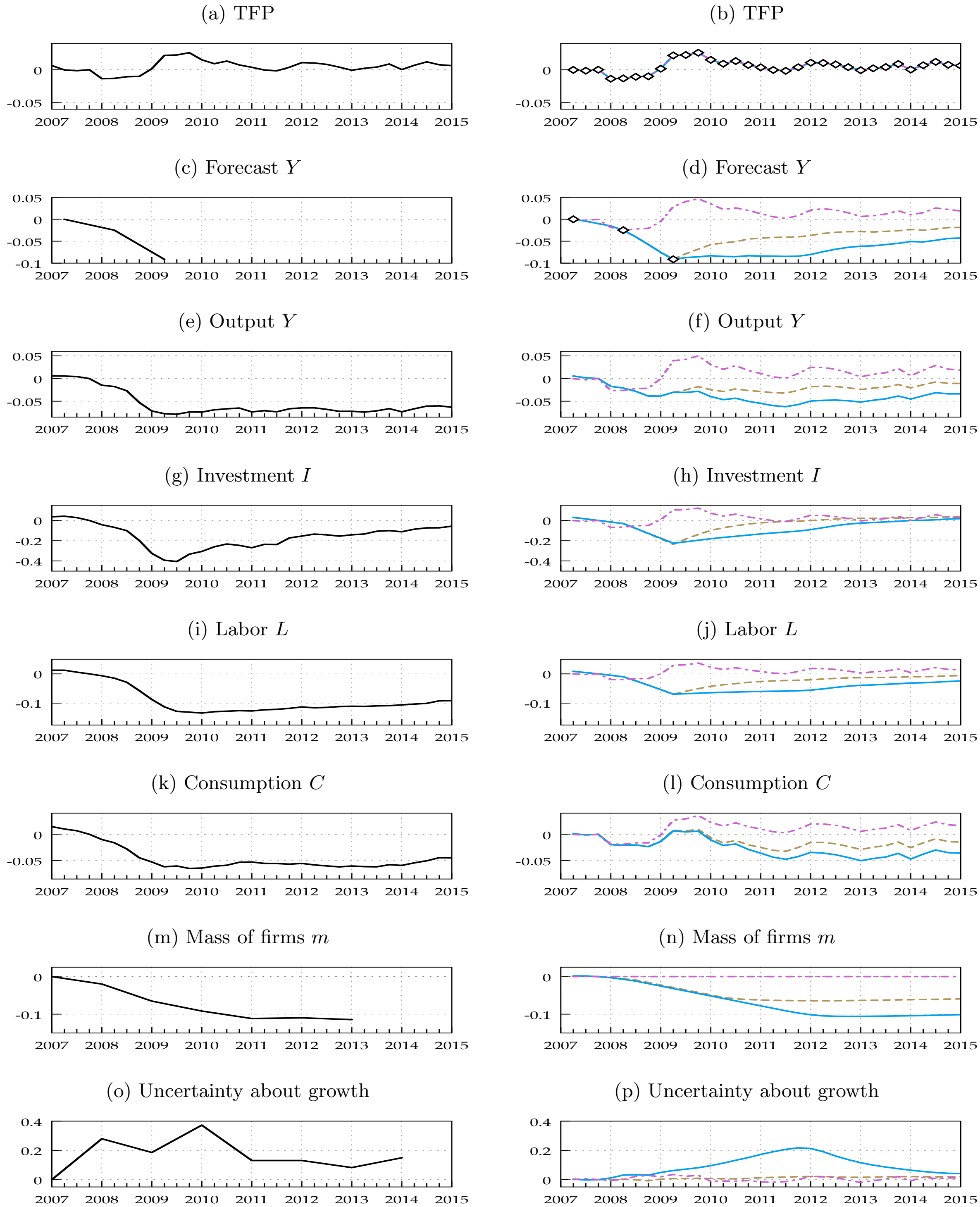
	Full	Fixed θ-uncertainty	RBC
A. Depth			
Benchmark	-7.8%	-6.0%	-2.6%
Full TFP fit	-6.2%	-3.8%	-2.6%
Linear detrending	-7.3%	-6.2%	-2.9%
PWT TFP	-8.0%	-7.2%	-5.5%
BLS TFP	-7.3%	-6.2%	-3.5%
CRRA Preferences	-3.9%	-3.9%	-1.9%
B. 20% Recovery			
Benchmark	2015:Q1	2012:Q4	2009:Q3
Full TFP fit	2012:Q1	2009:Q4	2009:Q1
Linear detrending	>2015:Q1	2013:Q3	2013:Q3
PWT TFP	>2015:Q1	>2015:Q1	2014:Q1
BLS TFP	>2015:Q1	2015:Q1	2014:Q2
CRRA Preferences	2012:Q2	2010:Q3	2008:Q3

Notes: We compute the depth and the duration of the recession from its official beginning in 2007:Q4 as determined by the NBER. The depth of the recession corresponds to the fall in output and the duration to the number of quarters until output reaches its lowest value.

Table 7: The depth and duration of the Great Recession as predicted by the three models in each of robustness exercise

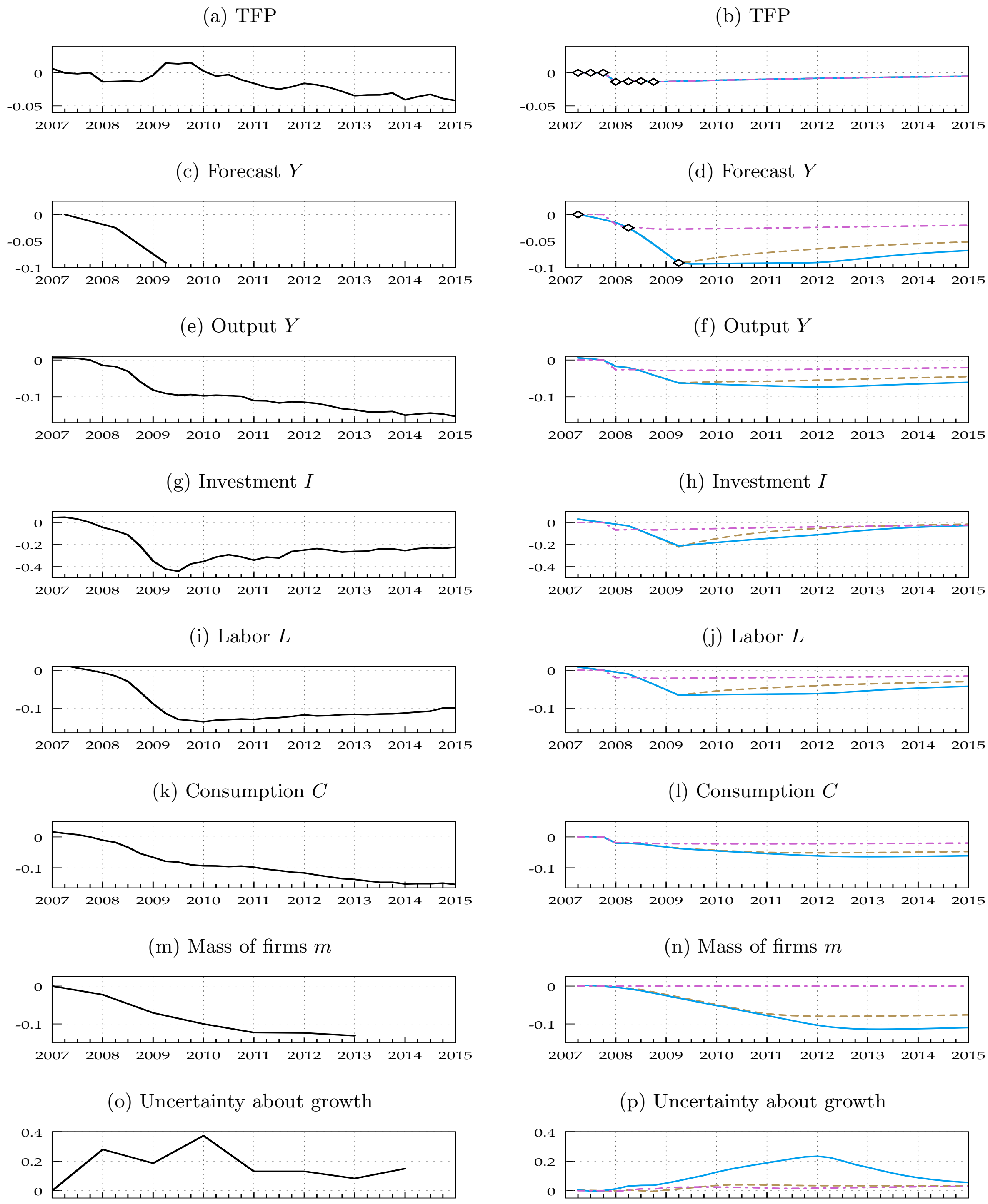
series of TFP instead of only the NBER recession period 2007-2009. In Figure 14, we use a standard linear detrending on the data instead of relying on a linear detrending with a structural break. In Figure 15, we use the TFP data from the Penn World Table instead of the capacity adjusted measure of Fernald. In Figure 16, we use the TFP data from the Bureau of Labor Statistics instead of the capacity adjusted measure of Fernald. In Figure 17, we use standard CRRA preferences instead of the GHH preferences. In all cases, we find that our full model generates a deeper and longer recession than the fixed θ-uncertainty and RBC alternatives. Table 7 shows the depth and the duration of the recession in the three models under all alternative exercises. Table 5 shows the cumulative losses.

The last exercise makes clear the reasons that pushed us to adopt GHH preferences as our benchmark. As in Jaimovich and Rebelo (2009), negative signals have peculiar implications in a model with wealth effects on the labor supply. Here, as the agents receive bad signals about the fundamental, they supply more labor relative to the economy with GHH preferences, and cut down on investment. As a result, consumption counterfactually rises for a short period of time. Other aggregates behaves roughly as in the data. In particular, our full model still explains a deeper and longer recession in output than the fixed θ-uncertainty and RBC alternatives.



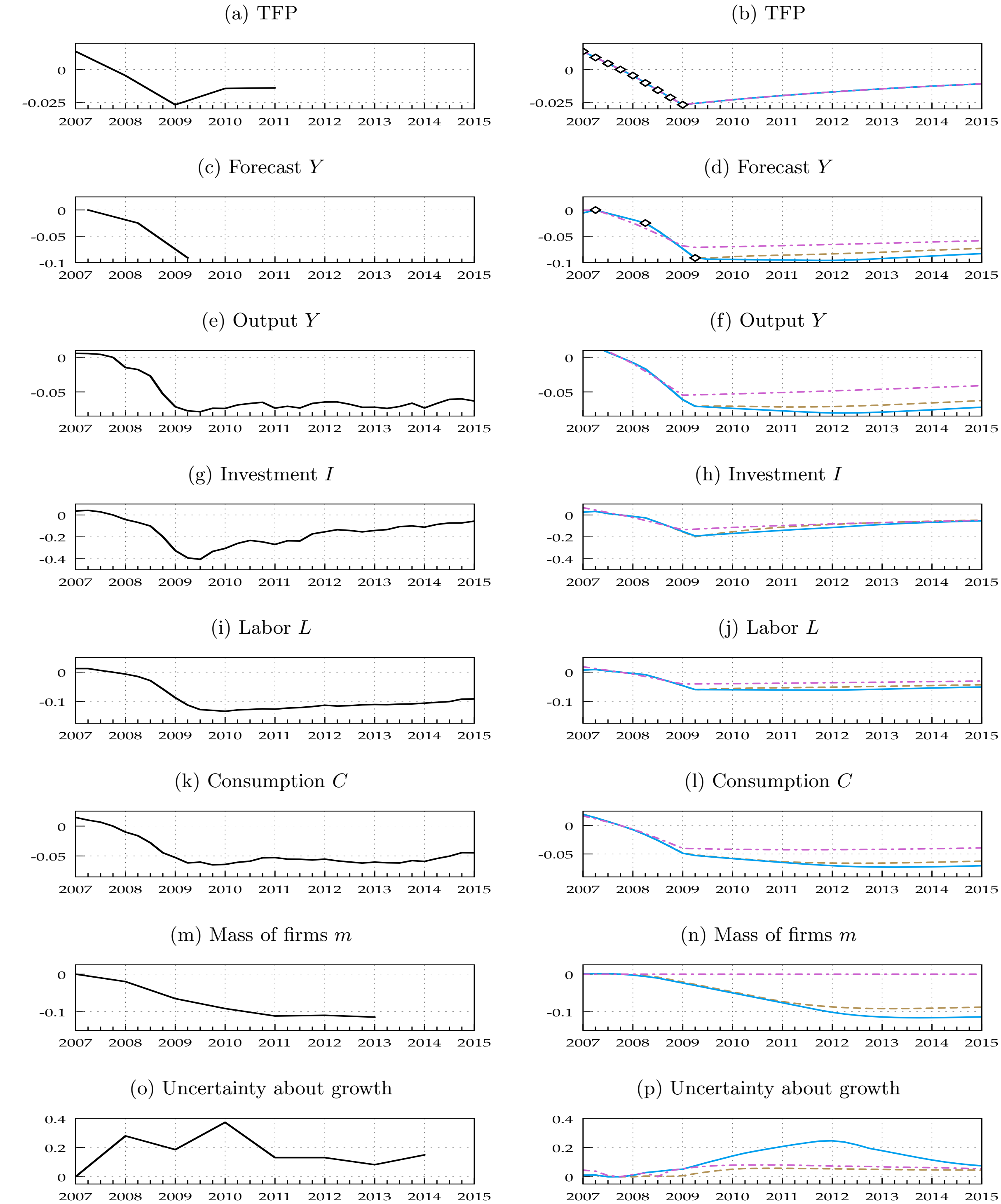
Notes: The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

Figure 13: Great Recession - Fitting the full time series of TFP



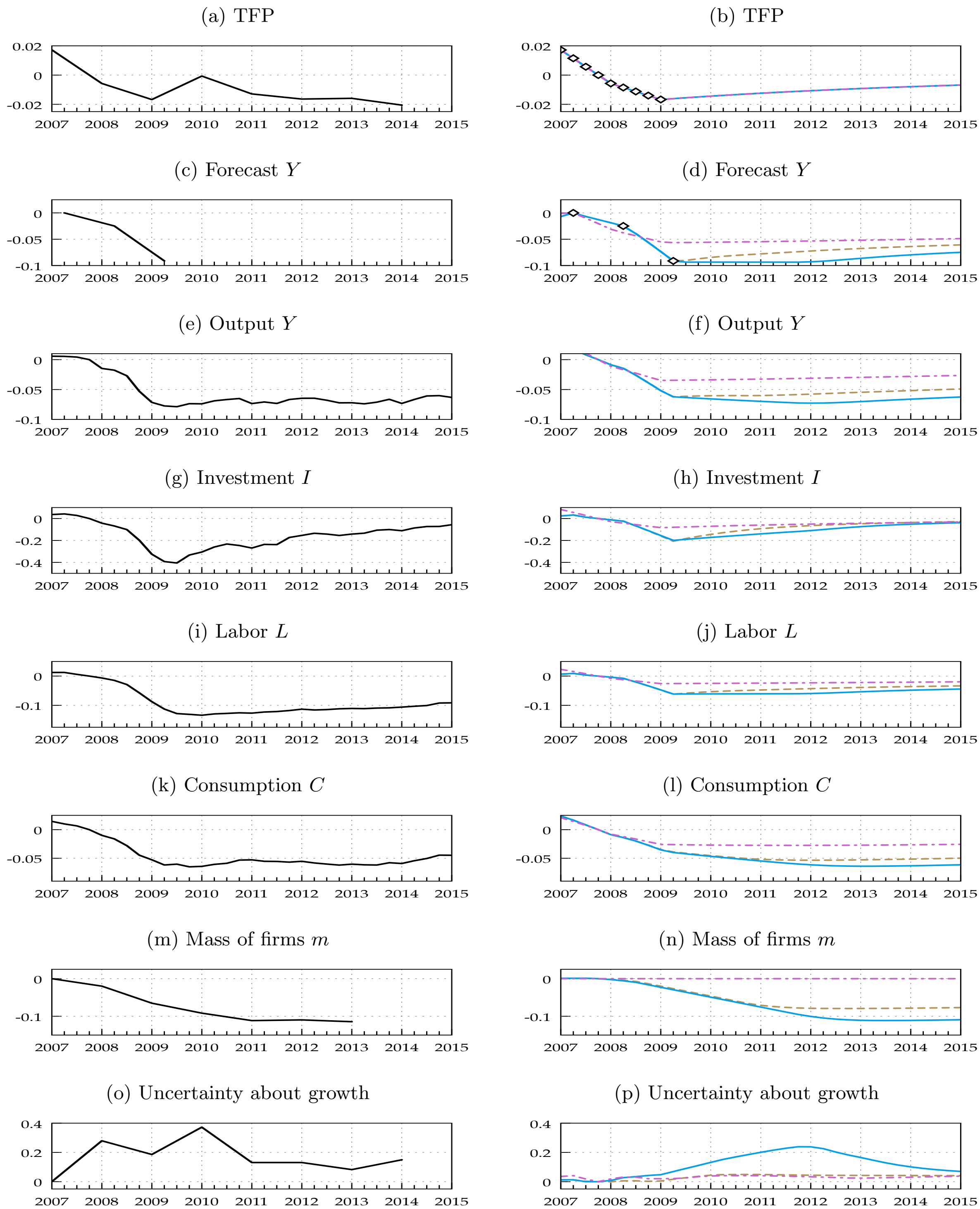
Notes: The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

Figure 14: Great Recession - linear detrending over 1978:Q1 to 2007:Q3



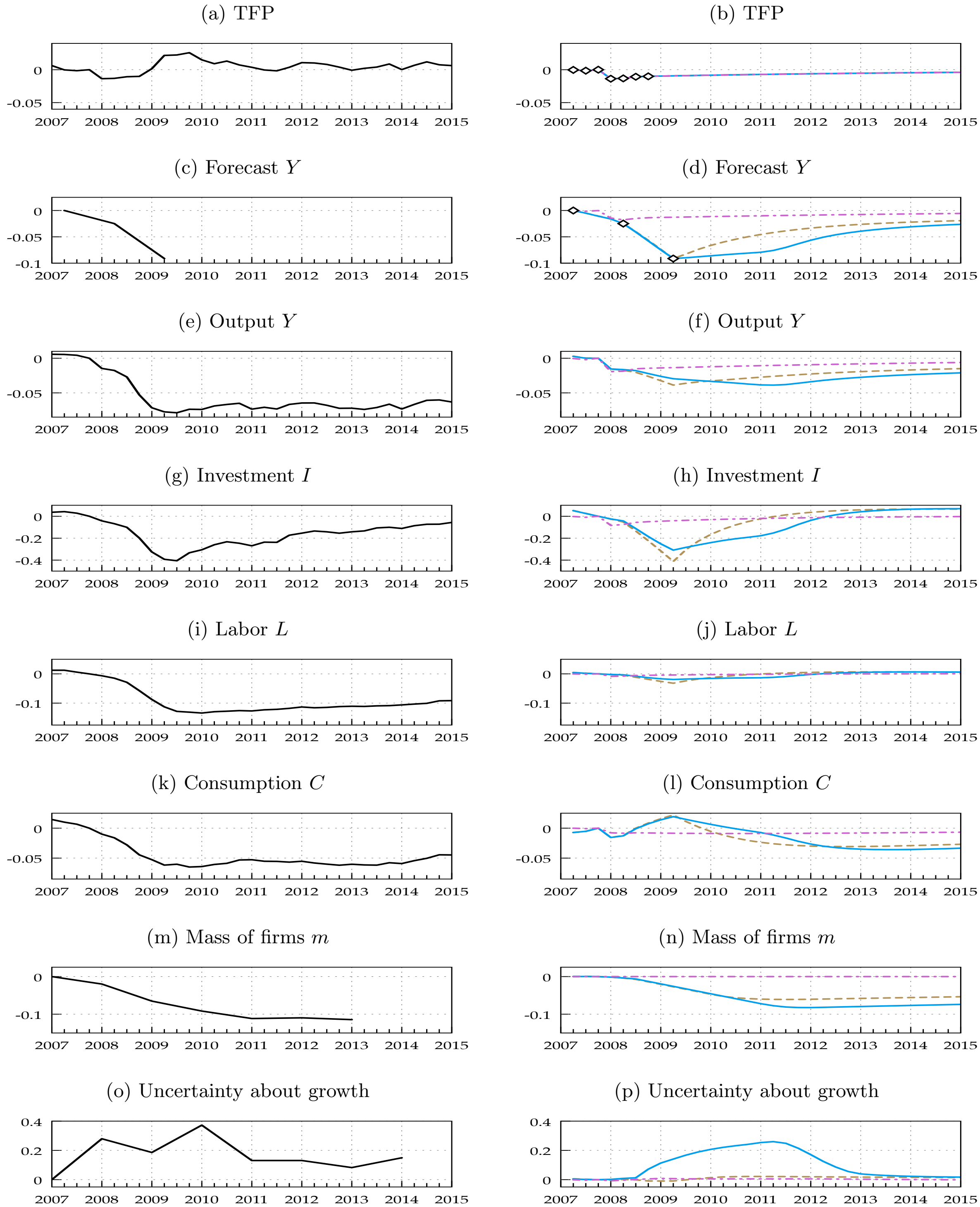
Notes: The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

Figure 15: Great Recession - TFP data from the Penn World Table



Notes: The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

Figure 16: Great Recession - TFP data from the Bureau of Labor Statistics

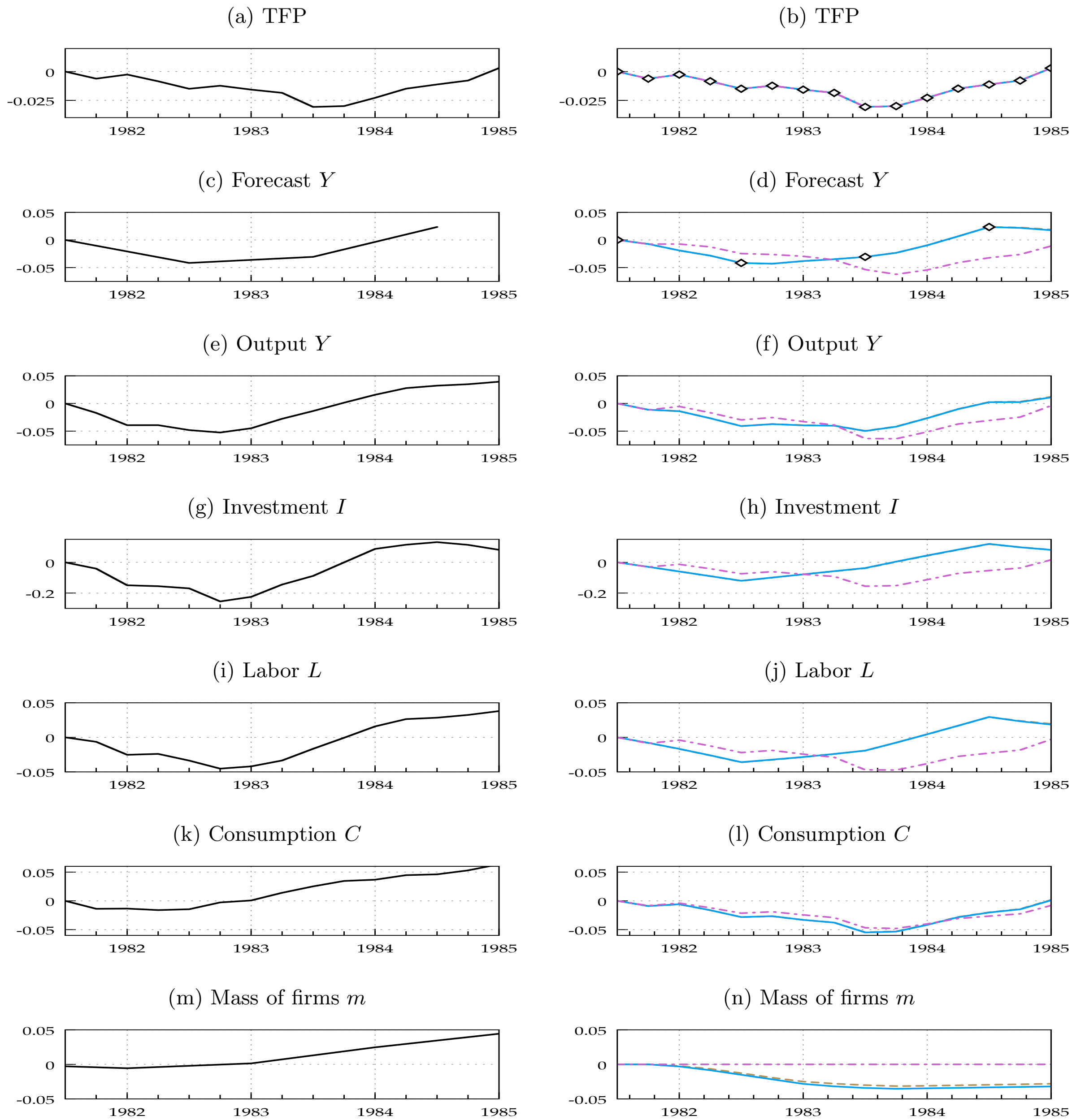


Notes: The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

Figure 17: Great Recession - CRRA preferences

D The 1981-1982 Recession

Figure 18 repeats the Great Recession exercise for the 1981-1982 recession. We do not include the behavior of uncertainty about real output growth as that period is not covered by the SPF data.[41]



Notes: The left column shows the data while the right column shows the models. In the right column, the solid curves show the full model, the dashed brown curves show fixed θ-uncertainty model and the dot-dash pink curves show the RBC model. All figures are in log deviation from trend. The scales are the same in both columns.

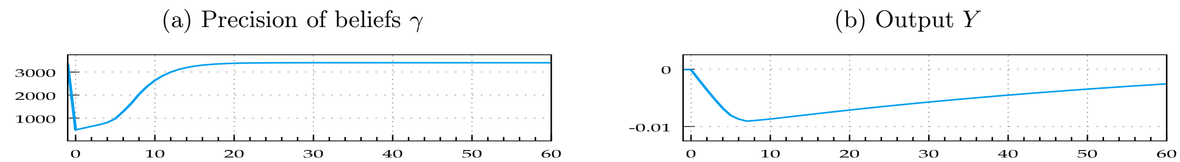
Figure 18: 1981-1982 Recession

[41] The SPF does have data about GNP growth over that period but a change from nominal to real GNP growth in the middle of the sample (1981Q3) prevents us from using this series.

E Further Applications

E.1 Uncertainty shocks

Figure 19 displays the impact of an uncertainty shock on the economy. We model this shock as an exogenous zero-probability event that raises uncertainty and that agents do not anticipate. The shock corresponds to a 100% increase in the standard deviation of the beliefs about the growth of output one year ahead.[42]

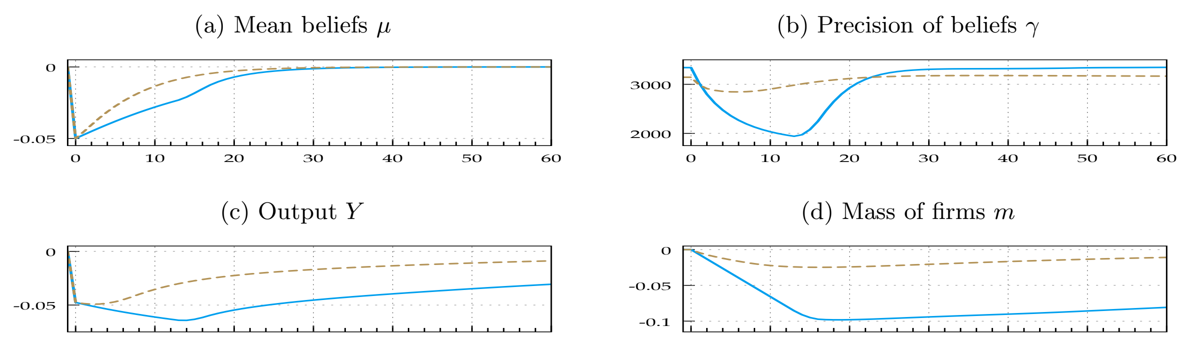


Notes: Evolution of the economy in the full model after an exogenous 100% increase in the standard deviation of the beliefs about the output growth one year ahead. Figure (a) is in level while Figure (b) is in log deviation from trend.

Figure 19: Impact of an exogenous uncertainty shock

E.2 Social planner

Figure 20 shows the response of the economy to a 5% negative shock to μ in the competitive equilibrium and in the planner's allocation.



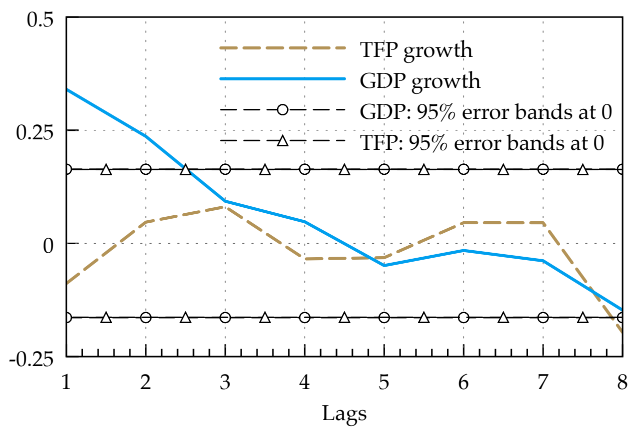
Notes: The solid curves show the evolution of the economy according to competitive equilibrium of the full model, while the dashed curves show the evolution of the economy under the social planner's allocation. Figures (c) and (d) are in log deviation from trend, the other figures are in level.

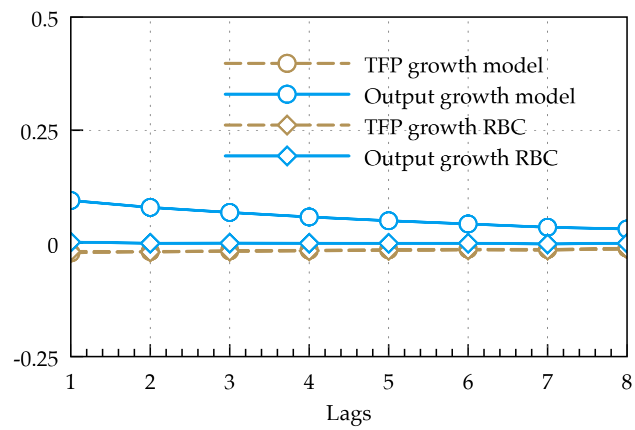
Figure 20: Response of the competitive equilibrium and the planner's allocation to a one period -5% shock to μ.

[42] Uncertainty shocks of this magnitude is not uncommon in the uncertainty-driven business cycle literature. See Bloom et al. (2012).

E.3 Autocorrelograms of output and TFP growth

Our benchmark model improves on the RBC model in terms of the propagation of shocks. As noted by Cogley and Nason (1995b), the RBC model features weak persistence in GDP growth. We repeat their exercise in Figure 21. Panel (a), showing the autocorrelograms of output and TFP growth in US data, implies that output growth is more persistent than TFP growth for smaller lags. The difference is also statistically significant at the 95% level. Panel (b) displays the same autocorrelograms in data generated by the RBC model and by our full model; because of the very long simulations, the error bands from model-generated autocorrelograms are essentially zero. Similarly to Cogley and Nason (1995b), the autocorrelogram of output growth in the RBC model mirrors that of TFP growth. In our full model, however, the persistence in output growth is larger than in TFP growth, and more so at lower lags. Hence, even though our model replicates only a fraction of the autocorrelation observed in the data, it does reproduce the *qualitative* feature of the data that the autocorrelation of GDP growth is larger than that of TFP growth at various lags. We interpret this finding as evidence that endogenous uncertainty leads to a more persistent series for output growth relative to TFP growth, in line with the baseline theory.



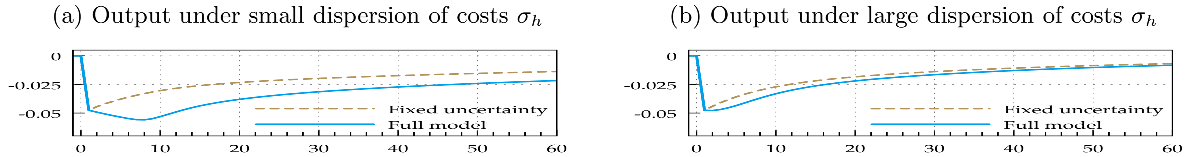


Notes. (a) The solid line is the autocorrelogram of the log GDP growth. We use the seasonally adjusted real GDP series from 1978:Q1 to 2015:Q1 from the Bureau of Economic Analysis after linearly detrending with a structural break in 2008:Q2 (see Appendix (A)). The dashed line is the autocorrelogram of the log of TFP growth using the capacity-adjusted measure from Fernald (2014), similarly detrended. The error bands show the 95% confidence interval at 0; (b) For comparison, both the RBC and the benchmark models are simulated with θ shocks only.

Figure 21: Autocorrelogram of output and TFP

E.4 Heterogeneous cost of entry

We consider the impact of heterogeneous costs of entry on the dynamics of the model. To do so, we assume that firms must pay a cost $f_h \sim \mathcal{N}\left(0, \sigma_h^2\right)$ when they enter. As a result, the entry decision is characterized by an equilibrium threshold such that only firms with fixed costs lower than the threshold enter. Figure 22 shows how changes in the dispersion of these fixed costs affect the response of the economy to a -5% shock to μ. When the dispersion of these shocks is relatively small, as in the left panel, the economy behaves similarly to the benchmark model. However, when the dispersion gets bigger, as in the right panel, the importance of aggregate uncertainty as a determinant of firm entry is reduced, as firms pay more attention to their idiosyncratic cost shock. In this case, the endogenous uncertainty about aggregate productivity is less important.



Notes: Evolution of output in the full model and the fixed θ-uncertainty model after a -5% shock to μ. The small shock is such that $\sigma_h/\bar{f} = 2\%$ while the large shock is such that $\sigma_h/\bar{f} = 4\%$.

Figure 22: Impact of heterogenous costs of entry

F Proofs

This section presents the proofs of the various lemmas and propositions stated in the paper. These proofs consider the limit version of the model as $\overline{N} \to \infty$, so that the fraction of investing firms is deterministic, i.e., $n = F(f^c)$.[43]

F.1 Lemmas 1 and 2

We start with the proofs of Lemmas 1 and 2, which characterize the laws governing the evolution of beliefs.

Lemma 1 (Full)**.** *For a given fraction of investing firms n, the mean of beliefs μ follows an AR(1) process with time-varying volatility s,*

$$\mu^{'} = \rho_\theta \mu + s(n,\gamma)\,\varepsilon,$$

where $s(n,\gamma) = \rho_\theta \left(\frac{1}{\gamma} - \frac{1}{\gamma+\gamma_z+n\gamma_x}\right)^{\frac{1}{2}}$ and $\varepsilon \sim \mathcal{N}(0,1)$. In addition, $s_n \geqslant 0$, $s_\gamma \leqslant 0$, $s_{nn} \leqslant 0$, $s_{\gamma\gamma} \leqslant 0$, while the sign of $s_{n\gamma}(n,\gamma)$ is positive for $\gamma \leqslant \frac{1}{2}(\gamma_z + n\gamma_x)$ and negative otherwise.

Proof. We use (2), (3) and (4) to compute the mean and the variance of the next period mean beliefs μ' given current-period information (μ,γ), and the fraction of investing firms n,

$$\mathbb{E}\left[\mu'|\mu,\gamma,n\right] = \rho_\theta \mu,$$
$$V\left[\mu'|\mu,\gamma,n\right] = \rho_\theta^2 \left(\frac{1}{\gamma} - \frac{1}{\gamma+\gamma_z+n\gamma_x}\right).$$

Being the sum of normally distributed variables, μ' is also normally distributed and can therefore be expressed as $\mu^{'} = \rho_\theta\mu + \rho_\theta\left(\frac{1}{\gamma} - \frac{1}{\gamma+\gamma_z+n\gamma_x}\right)^{\frac{1}{2}}\varepsilon$ with $\varepsilon \sim \mathcal{N}(0,1)$. The sign of the first and second

[43] Proofs in the finite $\overline{N}$ case are available upon request.

derivatives of $s(n,\gamma)$ are derived in Lemma 3 in the Online Appendix. In particular,

$$\begin{aligned} s_n(n,\gamma) &= \frac{\rho_\theta\gamma_x}{2(\gamma_z+n\gamma_x)^{\frac{1}{2}}}\frac{\gamma^{\frac{1}{2}}}{(\gamma+\gamma_z+n\gamma_x)^{\frac{3}{2}}} \geqslant 0, \\ s_\gamma(n,\gamma) &= -\frac{\rho_\theta(\gamma_z+n\gamma_z)^{\frac{1}{2}}}{2}\frac{2\gamma+\gamma_z+n\gamma_x}{\gamma^{\frac{3}{2}}(\gamma+\gamma_z+n\gamma_x)^{\frac{3}{2}}} \leqslant 0. \end{aligned}$$

□

Lemma 2 (Full). *The law of motion of the precision γ,*

$$\Gamma(n,\gamma) = \left(\frac{\rho_\theta^2}{\gamma+\gamma_z+n\gamma_x} + \left(1-\rho_\theta^2\right)\sigma_\theta^2\right)^{-1}, \tag{5}$$

satisfies the following:

1. *Γ is infinitely differentiable over $\mathbb{R}_+^2$ and increases with n and γ: $\Gamma_n \geqslant 0$ and $\Gamma_\gamma \geqslant 0$. In addition, $\Gamma_{nn} \leqslant 0$, $\Gamma_{n\gamma} \leqslant 0$ and $\Gamma_{\gamma\gamma} \leqslant 0$.*
2. *For a given n, there exists a unique positive fixed point $\gamma(n)$ in the precision of beliefs such that $\gamma(n) = \Gamma(n,\gamma(n))$. The function $\gamma(n)$ is strictly increasing in n.*

Proof. 1. The infinite differentiability follows from inspection of equation 5. The first and second derivatives of Γ are derived in Lemma 3 in the Online Appendix. In particular, we have

$$\Gamma_n(n,\gamma) = \frac{\rho_\theta^2\gamma_x}{(\gamma+\gamma_z+n\gamma_x)^2}\Gamma(n,\gamma)^2 \geqslant 0,$$

$$\Gamma_\gamma(n,\gamma) = \frac{\rho_\theta^2}{(\gamma+\gamma_z+n\gamma_x)^2}\Gamma(n,\gamma)^2 \geqslant 0,$$

as well as $\Gamma_{nn} \leqslant 0$, $\Gamma_{\gamma\gamma} \leqslant 0$, and $\Gamma_{n\gamma} \leqslant 0$.

2. Given n, existence and uniqueness of a positive fixed point follows from noting that all the fixed points γ must satisfy

$$0 = \gamma^2 + \gamma\left(\gamma_z + n\gamma_x - \frac{1}{\sigma_\theta^2}\right) - \frac{1}{\left(1-\rho_\theta^2\right)\sigma_\theta^2}(\gamma_z+n\gamma_x).$$

Because the quadratic function of γ on the right-hand side is convex in γ and negative at $\gamma = 0$, it necessarily has a unique positive root $\gamma(n)$. Differentiating the above equation with respect to n, we obtain

$$0 = 2\gamma(n)\gamma'(n) + \gamma'(n)\left(\gamma_z + n\gamma_x - \frac{1}{\sigma_\theta^2}\right) + \gamma_x\left(\gamma(n) - \frac{1}{\left(1-\rho_\theta^2\right)\sigma_\theta^2}\right),$$

which tells us that $\gamma'(n) = \frac{\gamma_x}{2\gamma(n)+\gamma_z+n\gamma_x-1/\sigma_\theta^2}\left(\frac{1}{(1-\rho_\theta^2)\sigma_\theta^2} - \gamma(n)\right)$. To see that $\gamma'(n) > 0$, note that

$$\frac{1}{\gamma(n)} = \frac{\rho_\theta^2}{\gamma(n)+\gamma_z+n\gamma_x} + \left(1-\rho_\theta^2\right)\sigma_\theta^2 > \left(1-\rho_\theta^2\right)\sigma_\theta^2$$

and

$$\frac{1}{\gamma(n)} = \frac{\rho_\theta^2}{\gamma(n)+\gamma_z+n\gamma_x} + \left(1-\rho_\theta^2\right)\sigma_\theta^2 < \frac{\rho_\theta^2}{\gamma(n)} + \left(1-\rho_\theta^2\right)\sigma_\theta^2,$$

which implies $\frac{1}{\sigma_\theta^2} < \gamma(n) < \frac{1}{(1-\rho_\theta^2)\sigma_\theta^2}$. □

F.2 Reformulation

To simplify the characterization of an equilibrium, we reformulate the problem of the firm in terms of a single mapping that characterizes the investment cutoff f^c. From the main text, the optimal investment strategy is to invest if and only if $f \leqslant f^c(\mu,\gamma)$ where

$$f^c(\mu,\gamma) = V^I(\mu,\gamma) - V^W(\mu,\gamma).$$

Substituting the value for $V^W(\mu,\gamma)$ from 7,

$$f^c(\mu,\gamma) \quad = \quad V^I(\mu,\gamma) - \beta\mathbb{E}\left\{\max\left[V^W\left(\mu',\gamma'\right), V^I\left(\mu',\gamma'\right) - f'\right]\right\}.$$

Subtracting the value $V^I(\mu',\gamma') - f'$ from the max operator and using the definition of $f^c(\mu,\gamma)$, we obtain that the following recursive mapping must be satisfied by the optimal cutoff

$$f^c(\mu,\gamma) = S(\mu,\gamma) - \beta\mathbb{E}\left\{\max\left[f' - f^c\left(\mu',\gamma'\right), 0\right]\right\}, \tag{25}$$

where $S(\mu,\gamma) = V^I(\mu,\gamma) - \beta\mathbb{E}\left[V^I(\mu',\gamma') - f'\right]$ is the opportunity cost of delaying investment for one period gross of the fixed cost f. In particular, using (4), (5) and the law of iterated expectations, we have

$$S(\mu,\gamma) \equiv \frac{1}{a}(1-\beta) + \beta\mu_f - \frac{1}{a}e^{-a\mu+\frac{a^2}{2\gamma}}\left(1 - \beta e^{a(1-\rho_\theta)\mu - \frac{a^2}{2}\frac{1-\rho_\theta^2}{\gamma} + \frac{a^2}{2}\left(1-\rho_\theta^2\right)\sigma_\theta^2}\right), \tag{26}$$

where μ_f is the mean of the fixed cost distribution, as introduced in the main text.

Remark. Note that the equilibrium cutoff function $f^c(\mu,\gamma)$ is enough to fully characterize the equilibrium. Given this function, we can recover all the other equilibrium objects. For all (μ,γ),

$$\begin{aligned} V^W(\mu,\gamma) &= V^I(\mu,\gamma) - f^c(\mu,\gamma), \\ V(\mu,\gamma) &= \beta\mathbb{E}\left\{V^I\left(\mu',\gamma'\right) - f' + \max\left[f' - f^c\left(\mu',\gamma'\right), 0\right]\right\}. \end{aligned}$$

whereas the fraction of investing firms is $n = F(f^c)$. Therefore, we focus on Bellman equation (25) for the optimal cutoff for the rest of the proofs.

F.3 Notation

In what follows, we denote $B(\mathcal{C})$ the set of bounded, continuous functions $g : \mathcal{C} \subseteq \mathbb{R}^l \longrightarrow \mathbb{R}$, for some $l \in \mathbb{N}$, equipped with the sup norm, $\|g\| = \sup_{x \in \mathcal{C}} |g(x)|$. We use $\overline{x}$ to denote an upper bound for some variable x, and $\underline{x}$ to denote a lower bound. For a given function $g \in B(\mathcal{C})$, we denote $g_{x_i}(x)$ the directional derivative $\frac{\partial g}{\partial x_i}(x) = \lim_{h \in \mathbb{R} \to 0} \frac{g(x+he_i)-g(x)}{h}$ where $e_i = (0, \ldots, 0, 1(i\text{th}), 0 \ldots)' \in \mathbb{R}^l$ when it exists. The notation $\overline{g}_{x_i}$and $\underline{g}_{x_i}$ is used to denote an upper or lower bound, respectively, on the ith directional derivative or a modulus of Lipschitz continuity whenever it applies.

F.4 Assumptions and Definitions

Before proving Propositions 1 to 4, we introduce a number of definitions and assumptions that will be used throughout the proofs. We first impose some regularity conditions on the distribution F of fixed costs f.

Assumption 1. *F is a continuously differentiable cumulative distribution function with bounded first derivatives. F has infinite support, mean μ_f and standard deviation σ_f, satisfies $F' > 0$ and $\int |f| \, dF < \infty$. We further assume that its mode $\text{argmax}_f F'(f)$ exists and is finite.*

We now define the set in which beliefs (μ, γ) lie. Note that because of the monotonicity in Γ and the existence of stationary points, the precision of beliefs γ remains bounded over time as long as the economy starts with a precision γ that belongs to $[\gamma(0), \gamma(1)]$. Indeed, if $\gamma \in [\gamma(0), \gamma(1)]$, then letting $\gamma' = \Gamma(n, \gamma)$, the properties of $\Gamma(n, \gamma)$ and $\gamma(n)$ from Lemma 2 imply that

$$\gamma(0) = \Gamma(0, \gamma(0)) \leqslant \Gamma(0, \gamma) \leqslant \gamma' \leqslant \Gamma(1, \gamma) \leqslant \Gamma(1, \gamma(1)) = \gamma(1).$$

We can thus restrict our attention to precisions belonging to the compact set $\left[\underline{\gamma}, \overline{\gamma}\right] = [\gamma(0), \gamma(1)]$.

The mean of beliefs μ, on the other hand, has an infinite support. In order to use the contraction mapping theorem in the space of bounded continuous functions, we restrict the space in which μ lies to a compact $\left[\underline{\mu}, \overline{\mu}\right]$. These bounds are chosen to be large enough as to contain almost all the ergodic distribution of μ. We thus introduce the following bounds:

Definition 4. *Define the following bounds and set:*

1. *Let $\overline{\gamma} \equiv \gamma(1)$ be the unique strictly positive solution to*

$$\overline{\gamma} = \left(\frac{\rho_\theta^2}{\overline{\gamma} + \gamma_z + \gamma_x} + \left(1 - \rho_\theta^2\right) \sigma_\theta^2 \right)^{-1} = \Gamma(1, \overline{\gamma}), \tag{27}$$

and $\underline{\gamma} \equiv \gamma(0) < \overline{\gamma}$ the unique strictly positive solution to

$$\underline{\gamma} = \left(\frac{\rho_\theta^2}{\underline{\gamma} + \gamma_z} + \left(1 - \rho_\theta^2\right) \sigma_\theta^2 \right)^{-1} = \Gamma\left(0, \underline{\gamma}\right). \tag{28}$$

2. *Let $\mathcal{S} = [\underline{\mu}, \overline{\mu}] \times \left[\underline{\gamma}, \overline{\gamma}\right]$, where $\underline{\mu}$ and $\overline{\mu}$ are some large bounds on μ.*

Because of the bounded support for μ, we must adapt the process for the evolution of the mean of beliefs. We do so by assuming that μ' follows a truncated Markov process.

Assumption 2. *The mean of beliefs μ' follows the truncated Markov process over $\left[\underline{\mu}, \overline{\mu}\right]$,*

$$\mu' = min\ \left\{max\left[\rho_\theta \mu + s\left(n,\gamma\right)\varepsilon, \underline{\mu}\right], \overline{\mu}\right\}, \tag{29}$$

where $s\left(n,\gamma\right)$ corresponds to the process in Lemma 1 and $\varepsilon \sim \mathcal{N}\left(0,1\right)$.

Notice that as the set $\left[\underline{\mu}, \overline{\mu}\right]$ becomes large, the truncated process μ' converges to the one described in Lemma 1.[44]

It is also useful for the rest of the proofs to introduce the following bounds on the derivatives of s and Γ.

Definition 5. *Define the following upper and lower bounds derived in Lemma 3 in the Online Appendix:*

1. *The derivatives of s are bounded by*

$$\begin{aligned}
\|s_\gamma\| &= \sup_{n\in[0,1],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]} \left|s_\gamma\left(n,\gamma\right)\right| = \left|s_\gamma\left(\gamma_x^{-1}\left(2\underline{\gamma}-\gamma_z\right),\underline{\gamma}\right)\right| = \rho_\theta\left(\frac{2}{3\underline{\gamma}}\right)^{\frac{3}{2}}, \\
\|s_n\| &= \sup_{n\in[0,1],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]} \left|s_n\left(n,\gamma\right)\right| = s_n\left(0,\frac{\gamma_z}{2}\right) = \frac{\rho_\theta}{\left(3\gamma_z\right)^{\frac{3}{2}}}\gamma_x;
\end{aligned}$$

2. *The derivatives of Γ are bounded above and below by*

$$\begin{aligned}
\overline{\Gamma}_\gamma &= \sup_{n\in[0,1],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]} \Gamma_\gamma\left(n,\gamma\right) = \rho_\theta^2\left(\frac{\underline{\gamma}}{\underline{\gamma}+\gamma_z}\right)^2, \\
\underline{\Gamma}_\gamma &= \inf_{n\in[0,1],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]} \Gamma_\gamma\left(n,\gamma\right) = \rho_\theta^2\left(\frac{\overline{\gamma}}{\overline{\gamma}+\gamma_z+\gamma_x}\right)^2, \\
\overline{\Gamma}_n &= \sup_{n\in[0,1],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]} \Gamma_n\left(n,\gamma\right) = \gamma_x\overline{\Gamma}_\gamma, \\
\underline{\Gamma}_n &= \inf_{n\in[0,1],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]} \Gamma_n\left(n,\gamma\right) = \gamma_x\underline{\Gamma}_\gamma.
\end{aligned}$$

To simplify our notation for the expectation operator under the truncated process (29), we define the following operator.

Definition 6. *Let R: $g \in B\left(\mathcal{S}\right) \to B\left(\mathbb{R}\times\left[\underline{\gamma},\overline{\gamma}\right]\right)$ be the operator that extends function g to the*

[44] In section F.2, we used the formula for the mean of a log-normal distribution to derive the expression of S. In the context of our truncated process μ, this formula should be understood as an approximation of arbitrary precision as $\left[\underline{\mu}, \overline{\mu}\right]$ becomes arbitrarily large.

domain $\mathbb{R} \times \left[\underline{\gamma}, \overline{\gamma}\right]$ *such that, for* $(\mu, \gamma) \in \mathbb{R} \times \left[\underline{\gamma}, \overline{\gamma}\right]$*:*

$$[R(g)](\mu,\gamma) = \begin{cases} g(\overline{\mu},\gamma) & \text{if } \mu > \overline{\mu}, \\ g(\mu,\gamma) & \text{if } \overline{\mu} \geq \mu \geq \underline{\mu}, \\ g(\underline{\mu},\gamma) & \text{if } \mu < \underline{\mu}. \end{cases}$$

Remark. It is equivalent to take the expectation of some function $g \in B(\mathcal{S})$ under the truncated process for μ in (29) or to take the expectation of the function $R(g) \in B\left(\mathbb{R} \times \left[\underline{\gamma}, \overline{\gamma}\right]\right)$ under the unrestricted process of (4). I.e., with Φ the CDF of a standard normal, we have:

$$\begin{aligned} \mathbb{E}[g \mid \mu, \gamma, n] &= \int_{\left(\underline{\mu}-\rho_\theta \mu\right)/s(n,\gamma)}^{(\overline{\mu}-\rho_\theta \mu)/s(n,\gamma)} g\left(\rho_\theta \mu + s(n,\gamma)\varepsilon, \Gamma(n,\gamma)\right)\Phi(d\varepsilon) \\ &\quad + \Phi\left(\frac{\underline{\mu}-\rho_\theta\mu}{s(n,\gamma)}\right) g\left(\underline{\mu}, \Gamma(n,\gamma)\right) + \left(1 - \Phi\left(\frac{\overline{\mu}-\rho_\theta\mu}{s(n,\gamma)}\right)\right) g\left(\overline{\mu}, \Gamma(n,\gamma)\right) \\ &= \int_{-\infty}^{\infty} R(g)\left(\rho_\theta\mu + s(n,\gamma)\varepsilon, \Gamma(n,\gamma)\right)\Phi(d\varepsilon). \end{aligned}$$

We now define the mapping $\mathcal{T}$ that characterizes the equilibrium investment cutoff as derived in equation (25).

Definition 7. *Let* $\mathcal{T} : f^c \in B(\mathcal{S}) \longrightarrow B(\mathcal{S})$ *be the mapping such that, for all* $(\mu, \gamma) \in \mathcal{S}$,

$$[\mathcal{T} f^c](\mu,\gamma) = S(\mu,\gamma) - \beta \iint \max\left[f' - R(f^c)\left(\rho_\theta\mu + s(n(\mu,\gamma),\gamma)\varepsilon, \Gamma(n(\mu,\gamma),\gamma)\right), 0\right] F\left(df'\right)\Phi(d\varepsilon), \tag{30}$$

where Φ *is the CDF of a standard normal,* $n(\mu,\gamma) = F(f^c(\mu,\gamma))$*, and where* $S(\mu,\gamma)$ *is defined in equation* (26)*.*

Under some conditions, we establish later that the mapping $\mathcal{T}$ admits a fixed point f^c in the set $\mathcal{F}_0$ defined as follows.

Definition 8. *Let* $\mathcal{F}_0 \subset B(\mathcal{S})$ *be the set of bounded continuous functions* $f^c : (\mu,\gamma) \in \mathcal{S} \longrightarrow \mathbb{R}$ *such that* f^c *is bounded by* $\overline{f}^c$*, Lipschitz continuous of modulus* $\overline{f}^c_\mu$ *in* μ *and* $\overline{f}^c_\gamma$ *in* γ*, where*

$$\overline{f}^c = (1-\beta)^{-1}\left[\frac{1}{a}(1-\beta) + \beta\left(\mu_f + \int |f|\, dF\right) + \frac{1}{a}e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}\left(1 - \beta e^{a(1-\rho_\theta)\underline{\mu} - \frac{a^2}{2}\frac{1-\rho_\theta^2}{\underline{\gamma}} + \frac{a^2}{2}\left(1-\rho_\theta^2\right)\sigma_\theta^2}\right)\right],$$

and $\left(\overline{f}^c_\mu, \overline{f}^c_\gamma\right) \in \mathbb{R}^2_+$ *solve the system*[45]

$$\begin{cases} \overline{f}^c_\mu = \left|S_\mu\left(\underline{\mu},\underline{\gamma}\right)\right| + \beta\overline{f}^c_\mu + \beta\left\|F'\right\|\left(\left\|s_n\right\|\overline{f}^c_\mu + \overline{\Gamma}_n\overline{f}^c_\gamma\right)\overline{f}^c_\mu, \\ \overline{f}^c_\gamma = \left|S_\gamma\left(\underline{\mu},\underline{\gamma}\right)\right| + \beta\left\|s_\gamma\right\|\overline{\Delta_s f} + \beta\overline{f}^c_\gamma + \beta\left\|F'\right\|\left(\left\|s_n\right\|\overline{f}^c_\mu + \overline{\Gamma}_n\overline{f}^c_\gamma\right)\overline{f}^c_\gamma, \end{cases} \tag{31}$$

with

$$\begin{aligned} S_\mu\left(\underline{\mu},\underline{\gamma}\right) &= e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}\left(1-\beta\rho_\theta e^{a(1-\rho_\theta)\underline{\mu}-\frac{a^2}{2}\frac{1-\rho_\theta^2}{\underline{\gamma}}+\frac{a^2}{2}\left(1-\rho_\theta^2\right)\sigma_\theta^2}\right) \\ S_\gamma\left(\underline{\mu},\underline{\gamma}\right) &= \frac{a}{2\gamma^2}e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}\left(1-\beta\rho_\theta^2 e^{a(1-\rho_\theta)\underline{\mu}-\frac{a^2}{2}\frac{1-\rho_\theta^2}{\underline{\gamma}}+\frac{a^2}{2}\left(1-\rho_\theta^2\right)\sigma_\theta^2}\right), \end{aligned}$$

and $\overline{\Delta_s f} = \overline{f}^c\left(\frac{1}{s(0,\overline{\gamma})} + \frac{1}{s(0,\overline{\gamma})^3}\right)$.

Later on, under additional parametric restrictions, we will be able to restrict the set in which the equilibrium cutoff function lies further to a subset $\mathcal{F}_1 \subset \mathcal{F}_0$ defined as follows.

Definition 9. *Let $\mathcal{F}_1 \subset \mathcal{F}_0$ be the subset of functions $f^c \in \mathcal{F}_0$ such that f^c is weakly increasing in μ and γ.*

We finally impose the following condition on parameters throughout the proofs:

Assumption 3. *The parameters and bounds are chosen so that*

$$1 \geqslant \beta e^{a(1-\rho_\theta)\overline{\mu}-\frac{a^2}{2}\frac{1-\rho_\theta^2}{\overline{\gamma}}+\frac{a^2}{2}\left(1-\rho_\theta^2\right)\sigma_\theta^2}. \tag{32}$$

This assumption is a necessary condition to guarantee that the real option channel is active. The first two terms in the exponential require that the mean reversion in μ and γ is small enough not to dominate the wait-and-see effects. The last term requires that the variance of the fundamental σ_θ^2 is not too large to discourage firms from waiting. Note that Condition 32 is satisfied as $\rho_\theta \to 1$.

F.5 Proposition 1

Proposition 1 (Full)**.** *Under Assumptions 1-3 and for parameters such that*[46]

$$\frac{4\beta}{(1-\beta)^2}\left\|F'\right\|\left[\left\|s_n\right\|\left|S_\mu\left(\underline{\mu},\underline{\gamma}\right)\right| + \overline{\Gamma}_n\left|S_\gamma\left(\underline{\mu},\underline{\gamma}\right)\right| + \beta\overline{\Gamma}_n\left\|s_\gamma\right\|\overline{\Delta_s f}\right] \leqslant 1, \tag{33}$$

[45] Condition (33) in Proposition 1 is a sufficient condition that guarantees that this system of quadratic equations admits a strictly positive solution $\left(\overline{f}^c_\mu, \overline{f}^c_\gamma\right)$. In the case where there are multiple solutions (at most two), we pick the maximal one.

[46] Note that condition (33) is only a sufficient condition, so that an equilibrium may exist under weaker conditions. Since $\|s_n\| = O(\gamma_x)$ and $\overline{\Gamma}_n = O(\gamma_x)$ while the other quantities in the condition remain bounded as $\gamma_x \to 0$, condition (33) is always satisfied as long as γ_x is small enough, implying that the complementarity in information is weak.

where $\overline{\Delta_s f} = \overline{f}^c \left(\frac{1}{s(0,\overline{\gamma})} + \frac{1}{s(0,\overline{\gamma})^3}\right)$, *there exists a unique equilibrium such that the optimal cutoff* f^c *belongs to* $\mathcal{F}^0$. *If, in addition,*[47]

$$\begin{aligned} S_\mu(\overline{\mu},\overline{\gamma}) &\geqslant \beta\overline{\Delta_n f}\left\|F'\right\|\overline{f}^c_\mu, &(34)\\ S_\gamma(\overline{\mu},\overline{\gamma}) &\geqslant \beta\left(\overline{\Delta_s f}\left\|s_\gamma\right\| + \overline{\Delta_n f}\left\|F'\right\|\overline{f}^c_\gamma\right), &(35)\end{aligned}$$

where $\overline{\Delta_n f} = \|s_n\|\,\overline{f}^c_\mu + \overline{\Gamma}_n\overline{f}^c_\gamma$ *then the resulting optimal cutoff* $f^c(\mu,\gamma)$ *is increasing in* μ *and* γ.

Proof. *Outline:* The equilibrium mapping $\mathcal{T}$ that we consider corresponds to an inefficient competitive equilibrium subject to an informational externality. Because of the feedback of other investors' actions on individual beliefs, standard existence/uniqueness arguments do not directly apply. In particular, a strong enough feedback could potentially lead to equilibrium multiplicity. The magnitude of the feedback is, however, governed by the informativeness γ_x of the social learning channel. Our proof establishes that this feedback is a $O(\gamma_x)$. Hence, when γ_x is low, the complementarity in information is negligible and we prove that the mapping defined in Definition 7 is a contraction.

We proceed in four steps: 1. We show that $\mathcal{T}$ is a well-defined self-map on $\mathcal{F}_0$ under condition (33); 2. We establish that $\mathcal{F}_0$ is a Banach space; 3. We show that $\mathcal{T}$ is a contraction on $\mathcal{F}_0$ when γ_x is small enough; 4. Under the additional conditions on parameters (34) and (35), which imply that various feedback effects are dominated by the first order impact of μ and γ in the objective function S, we characterize further the monotonicity of the optimal cutoff.

Step 1. *$\mathcal{T}$ is a well-defined mapping from $\mathcal{F}_0$ to $\mathcal{F}_0$.* Under condition 33, Lemma 5 in the Online Appendix shows that there exists a solution to the system of equations (31) which defines the moduli of Lipschitz continuity and Lemma 6 in the Online Appendix demonstrates, under the same conditions, that $\mathcal{T}$ maps $\mathcal{F}_0$ onto itself. $\mathcal{T}$ is thus a well-defined self-map on $\mathcal{F}^0$.

Step 2. *$\mathcal{F}_0$ is a Banach space.* $\mathcal{F}_0$ is a subset of the Banach space of bounded continuous functions $B(\mathcal{S})$ equipped with the sup norm $\|\cdot\|$. We now show that $\mathcal{F}_0$ is closed. Take a Cauchy sequence $\{f^c_k\}_{k\in\mathbb{N}} \in \prod_{k\in\mathbb{N}}\mathcal{F}^0$. First, since $B(\mathcal{S})$ is a Banach space, it converges to some $f^c \in B(\mathcal{S})$. Second, the bound $\overline{f}^c$ is preserved in the limit ($f^c_k(\mu,\gamma) \leqslant \overline{f}^c \to f^c(\mu,\gamma) \leqslant \overline{f}^c$). Third, the Lipschitz property is preserved in the limit. Indeed, we have

$$\forall(\mu_1,\mu_2), k\in\mathbb{N}, \left|f^c_k(\mu_2,\gamma) - f^c_k(\mu_1,\gamma)\right| \leqslant \overline{f}^c_\mu\left|\mu_2-\mu_1\right|.$$

Taking the limit $k\to\infty$, we then have $|f^c(\mu_2,\gamma) - f^c(\mu_1,\gamma)| \leqslant \overline{f}^c_\mu|\mu_2-\mu_1|$, so that f^c is also Lipschitz in μ with modulus $\overline{f}^c_\mu$. Applying the same argument along γ implies that the limit f^c is also Lipschitz in γ of modulus $\overline{f}^c_\gamma$. Hence, $\mathcal{F}^0$ is closed. Being a closed subset of a Banach space, $\mathcal{F}_0$ is a Banach space.

Step 3. *$\mathcal{T}$ is a contraction on $\mathcal{F}_0$* under condition 33. Let $f^c_1, f^c_2 \in \mathcal{F}_0$. Denote $n_i = F(f^c_i), i =$

[47] Conditions 34 and 35 are satisfied when, for instance, 1) β is small enough, or 2) risk aversion a is high enough that S_γ is large enough and γ_x is small enough that $\overline{\Delta_n f}$ is small.

1, 2. We decompose the differential as follows:

$$\mathcal{T} f_2^c (\mu, \gamma) - \mathcal{T} f_1^c (\mu, \gamma) \quad = \quad \beta (A + B)$$

where

$$\begin{aligned} A \quad &= \quad \iint \Big[\max \left(f' - [R (f_1^c)] \left(\rho_\theta \mu + s (n_1 (\mu, \gamma), \gamma) \varepsilon, \Gamma (n_1 (\mu, \gamma), \gamma) \right), 0 \right) \\ &\quad - \max \left(f' - [R (f_2^c)] \left(\rho_\theta \mu + s (n_1 (\mu, \gamma), \gamma) \varepsilon, \Gamma (n_1 (\mu, \gamma), \gamma) \right), 0 \right) \Big] F \left(df' \right) \Phi (d\varepsilon) \\ B \quad &= \quad \iint \Big[\max \left(f' - [R (f_2^c)] \left(\rho_\theta \mu + s (n_1 (\mu, \gamma), \gamma) \varepsilon, \Gamma (n_1 (\mu, \gamma), \gamma) \right), 0 \right) \\ &\quad - \max \left(f' - [R (f_2^c)] \left(\rho_\theta \mu + s (n_2 (\mu, \gamma), \gamma) \varepsilon, \Gamma (n_2 (\mu, \gamma), \gamma) \right), 0 \right) \Big] F \left(df' \right) \Phi (d\varepsilon) \end{aligned}$$

We now provide bounds for each term separately. Regarding term A, using Lemma 4.1, we have

$$\begin{aligned} |A| \quad &\leqslant \quad \int \Big| R (f_2^c) \left(\rho_\theta \mu + s (n_1 (\mu, \gamma), \gamma) \varepsilon, \Gamma (n_1 (\mu, \gamma), \gamma) \right) \\ &\quad - R (f_1^c) \left(\rho_\theta \mu + s (n_1 (\mu, \gamma), \gamma) \varepsilon, \Gamma (n_1 (\mu, \gamma), \gamma) \right) \Big| \Phi (d\varepsilon) \\ &\leqslant \quad \| f_2^c - f_1^c \| . \end{aligned}$$

Turning to term B, using the Lipschitz property along μ and γ, and since the Lipschitz property is preserved under operator R (Lemma 4.3), the first part of Lemma 8 tells us that

$$\begin{aligned} |B| \quad &= \quad \overline{\Delta_n f} \left| F (f_2^c (\mu, \gamma)) - F (f_1^c (\mu, \gamma)) \right| \\ &\leqslant \quad \overline{\Delta_n f} \left\| F' \right\| \| f_2^c - f_1^c \| . \end{aligned}$$

where $\overline{\Delta_n f} = \| s_n \| \, \overline{f}_\mu^c + \overline{\Gamma}_n \overline{f}_\gamma^c$. Note that $\overline{\Delta_n f}$ captures the feedback effect that changes in n have on expectations of future investment returns. From the expressions in Definition 5 we have that $\| s_n \| \to 0$, $\overline{\Gamma}_n \to 0$, $\overline{f}_\mu^c \to (1 - \beta)^{-1} \left| S_\mu \left(\underline{\mu}, \underline{\gamma} \right) \right|$, and $\overline{f}_\gamma^c \to (1 - \beta)^{-1} \left[\left| S_\gamma \left(\underline{\mu}, \underline{\gamma} \right) \right| + \beta \| s_\gamma \| \, \overline{\Delta_s f} \right]$ as $\gamma_x \to 0$, hence $\overline{\Delta_n f}$ is $O(\gamma_x)$. This reflects that changes in the expectation of n only affect information through the precision of the social information channel γ_x. Hence, B is negligible for γ_x sufficiently small. As a result, the mapping $\mathcal{T}$ is a contraction if

$$\beta \left(1 + \left\| F' \right\| \overline{\Delta_n f} \right) < 1,$$

which is satisfied under the same condition (33) guaranteeing that the system (31) has a solution (see Lemma 5). Indeed, the first row of system (31) tells us that

$$\beta \left(1 + \left\| F' \right\| \overline{\Delta_n f} \right) \overline{f}_\mu^c = \overline{f}_\mu^c - \left| S_\mu \left(\underline{\mu}, \underline{\gamma} \right) \right| .$$

Hence, we have $\beta\left(1+\|F'\|\,\overline{\Delta_n f}\right)=1-\left|S_\mu\left(\underline{\mu},\underline{\gamma}\right)\right|/\overline{f}^c_\mu<1$. $\mathcal{T}$ is thus a contraction under condition (33), so that there exists a unique equilibrium such that $f^c\in\mathcal{F}_0$.

Step 4. Monotonicity of the cutoff f^c. If, in addition, conditions (34) and (35) are satisfied, then Lemma 7 applies and $\mathcal{T}$ maps $\mathcal{F}_1$ onto itself. Because the weak monotonicity properties are preserved in the limit, $\mathcal{F}_1$ is closed. Because $\mathcal{T}$ is also a contraction on $\mathcal{F}_1\subset\mathcal{F}_0$, we conclude that the equilibrium cutoff f^c belongs to $\mathcal{F}_1$. In particular, it is weakly increasing in μ and γ. □

F.6 Proposition 2

Proposition 2 (Full)**.** *Under Assumptions 1-3 and conditions* (33)*,* (34) *and* (35)*, if parameters are such that*[48]

$$\frac{1}{2-\beta}\left\|F'\right\|\left[S_\gamma\left(\overline{\mu},\overline{\gamma}\right)-\beta\overline{\Delta_s f}\left\|s_\gamma\right\|\right]\underline{\Gamma_n}+\underline{\Gamma_\gamma}>1, \tag{36}$$

where $\underline{\Gamma_\gamma}=\rho_\theta^2\left(\frac{\overline{\gamma}}{\overline{\gamma}+\gamma_z+\gamma_x}\right)^2$, $\underline{\Gamma_n}=\gamma_x\underline{\Gamma_\gamma}$ *and* $\overline{\Delta_s f}=\overline{f}^c\left(\frac{1}{s(0,\overline{\gamma})}+\frac{1}{s(0,\overline{\gamma})^3}\right)$*, then the economy admits an uncertainty trap. In particular, there exists a non-empty interval* (μ_l,μ_h) *such that, for all* $\mu\in(\mu_l,\mu_h)$*, the economy features at least two stationary points* $\gamma_l(\mu)<\gamma_h(\mu)$ *in the dynamics of belief precision. Regime* γ_l *is characterized by high uncertainty and low investment while regime* γ_h *is characterized by low uncertainty and high investment.*

Proof. Outline: The purpose of this proof is to establish that there may exist various stationary points in the dynamics of γ for a range of μ's. It is therefore useful to introduce the following function:

$$\varphi(n,\gamma)\equiv\Gamma(n,\gamma)-\gamma=\left(\frac{\rho_\theta^2}{\gamma+\gamma_z+n\gamma_x}+\left(1-\rho_\theta^2\right)\sigma_\theta^2\right)^{-1}-\gamma, \tag{37}$$

where Γ is the law of motion for γ defined in (5). By definition, stationary points in the dynamics of γ satisfy $\varphi(n(\mu,\gamma),\gamma)=0$. We show that there exists a non-empty set (μ_l,μ_h) such that, for all $\mu\in(\mu_l,\mu_h)$, there exist $\gamma_l(\mu)<\gamma_h(\mu)$ such that $\varphi(n(\mu,\gamma_l(\mu)),\gamma_l(\mu))=\varphi(n(\mu,\gamma_h(\mu)),\gamma_h(\mu))=0$, i.e., $\gamma_l(\mu)$ and $\gamma_h(\mu)$ are two distinct stationary points in the dynamics of γ for some given mean beliefs μ.

We proceed in two steps. In step 1, we show that we can find a stationary point (μ^*,γ^*) with $\varphi(n(\mu^*,\gamma^*),\gamma^*)=0$ such that the function $\gamma\mapsto\varphi(n(\mu^*,\gamma),\gamma)$ crosses the x-axis at γ^* from below. Note that this condition captures the situation depicted on Figure 5, where the law of motion for γ crosses the 45° line from below at least once. To show this, we establish that the equilibrium cutoff f^c is increasing with γ at a strictly positive rate bounded away from 0 under condition (36). Hence, when the distribution of fixed cost F is sufficiently concentrated, the number of firms switching from waiting to investing, after an increase in γ, becomes large enough that φ crosses the x-axis

[48] Condition (36) is satisfied in particular if either 1) the dispersion of the fixed cost distribution σ_f is small enough that $\|F'\|$ is large enough, or 2) if the risk aversion parameter a is high enough that S_γ is large enough. Note also that the set of parameters satisfying conditions (33), (34), (35) and (36) is non-empty for 1) β small enough, or 2) for $\|F'\|$ and a large enough.

from below. In step 2, we show that this implies that at least two stationary points exist for μ^*, and that uncertainty traps must then also exist in a neighborhood of μ^*.

Step 1. The proof is constructive. We first establish that there exist a γ^* such that $\varphi(F(f_m),\gamma^*) = 0$, where $f_m \equiv \text{argmax } F'$ is the mode of F. Intuitively, φ is more likely to cross the x-axis from below at a point where a large number of firms switches from waiting to investing, i.e., when the cutoff f^c is close to the mode of the distribution of fixed costs. From the definition of $\{\underline{\gamma},\overline{\gamma}\}$ in (27) and (28) and the monotonicity of Γ in n, we have $\varphi(n,\underline{\gamma}) \geq 0 \geq \varphi(n,\overline{\gamma})$ for all $n \in [0,1]$, in particular for $n = F(f_m)$. Since Γ is continuous in n and γ, φ is also continuous, and there exists $\gamma^* \in [\underline{\gamma},\overline{\gamma}]$ such that $\varphi(F(f_m),\gamma^*) = 0$.

We then establish that there is μ^* such that $n(\mu^*,\gamma^*) = F(f_m)$. For γ_x small enough that the conditions (33), (34) and (35) are satisfied, we know that the equilibrium cutoff f^c belongs to $\mathcal{F}_1$. In particular, f^c is Lipschitz continuous and increasing in μ. In addition, Lemma 7 in the Online Appendix tells us that the rate at which f^c increases in μ is bounded from below and this bound is strictly positive for $\mu \in [\underline{\mu},\overline{\mu}]$ under condition (34). By continuity, this implies that there exists $\mu^* \in [\underline{\mu},\overline{\mu}]$ such that $f^c(\mu^*,\gamma^*) = f_m$ as long as the bounds $\underline{\mu}$ and $\overline{\mu}$ are chosen large enough. By construction, $\varphi(n(\mu^*,\gamma^*),\gamma^*) = 0$.

At this stage we have established that there is $(\mu^*,\gamma^*) \in \mathcal{S}$ corresponding to a fixed point in the dynamics of γ at which the investment cutoff corresponds to the mode of the distribution of fixed costs. It remains to be shown that $\varphi(n(\mu^*,\gamma),\gamma)$ crosses zero from below at $\gamma = \gamma^*$.

Let $\epsilon > 0$ be arbitrarily small. By continuity of F' and the bounds on the derivatives of Γ, there exists a neighborhood Ω of γ^* such that $\inf_{\gamma\in\Omega} F'(f^c(\mu^*,\gamma)) \geqslant (1-\epsilon)\|F'\|$. Using the bound provided in Lemma 9, we have that, for all $\gamma \in \Omega$,

$$\begin{aligned}
\varphi(n(\mu^*,\gamma),\gamma) &\geqslant \underline{\varphi}_\gamma^\epsilon(\gamma-\gamma^*), \text{ for } \gamma \geqslant \gamma^*,\\
\varphi(n(\mu^*,\gamma),\gamma) &\leqslant \underline{\varphi}_\gamma^\epsilon(\gamma-\gamma^*), \text{ for } \gamma \leqslant \gamma^*,
\end{aligned}$$

where $\underline{\varphi}_\gamma^\epsilon \equiv (1-\epsilon)\underline{\Gamma_n}\left(1+\beta\overline{f}_\mu^c\|F'\|\|s_n\|\right)^{-1}\|F'\|\left[S_\gamma(\overline{\mu},\overline{\gamma})-\beta\overline{\Delta_s f}\|s_\gamma\|\right]+\underline{\Gamma_\gamma}-1$. We now show that $\underline{\varphi}_\gamma^\epsilon > 0$ under conditions (33) and (36). Note first that $\overline{f}_\mu^c\|s_n\| \leqslant \overline{\Delta_n f}$, where $\overline{\Delta_n f}$ is defined in Lemma 8. We know from Proposition 1 that the necessary condition (33) for the existence of an equilibrium implies $\beta\left(1+\|F'\|\overline{\Delta_n f}\right) < 1$. Hence, $1+\beta\|F'\|\overline{\Delta_n f} \leqslant 2-\beta$ and we have

$$\varphi_\gamma^\epsilon \geqslant (1-\varepsilon)\frac{1}{2-\beta}\left\|F'\right\|\left[S_\gamma(\overline{\mu},\overline{\gamma})-\beta\overline{\Delta_s f}\|s_\gamma\|\right]\underline{\Gamma_n}+\underline{\Gamma}_\gamma-1,$$

which, under condition (36) and for ϵ small enough, implies $\varphi_\gamma^\epsilon > 0$. We have thus shown that the function $\varphi(n(\mu^*,\gamma),\gamma)$ crosses zero at γ^* from below.

Step 2. Since φ crosses the x-axis from below at γ^*, there exists some (γ_+,γ_-) in the neighborhood of γ^* such that $\gamma_- < \gamma^* < \gamma_+$ and $\varphi(n(\mu^*,\gamma_-),\gamma_-) < 0 < \varphi(n(\mu^*,\gamma_+),\gamma_+)$. By continuity of φ, there exists a neighborhood (μ_l,μ_h) of μ^* such that, for all $\mu \in (\mu_l,\mu_h)$, $\varphi(n(\mu,\gamma_-),\gamma_-) < 0 < \varphi(n(\mu,\gamma_+),\gamma_+)$. From the definition of $\{\underline{\gamma},\overline{\gamma}\}$ and the monotonicity of f^c, we have that, for all $\mu \in (\mu_l, \mu_h)$, $\varphi\left(n\left(\mu, \underline{\gamma}\right), \underline{\gamma}\right) \geqslant 0 \geqslant \varphi\left(n\left(\mu, \overline{\gamma}\right), \overline{\gamma}\right)$. Therefore, by the Intermediate Value theorem, there exists $\gamma_l(\mu) \in [\underline{\gamma}, \gamma_-)$ and $\gamma_h(\mu) \in (\gamma_+, \overline{\gamma}]$ such that $\varphi\left(n\left(\mu, \gamma_l(\mu)\right), \gamma_l(\mu)\right) = \varphi\left(n\left(\mu, \gamma_h(\mu)\right), \gamma_h(\mu)\right) = 0$. We have thus found two distinct stationary points in γ for all μ in a neighborhood (μ_l, μ_h) of μ^*. From the strict monotonicity of f^c in γ, $n\left(\mu, \gamma_l(\mu)\right) < n\left(\mu, \gamma_h(\mu)\right)$, which implies that the low regime $\gamma_l(\mu)$ features lower investment and higher uncertainty than the high regime $\gamma_h(\mu)$. □

F.7 Proposition 3

We now introduce the mapping satisfied by the constrained social planner's problem.

Definition 10. Let $\mathcal{T}^W : B(\mathcal{S}) \to B(\mathcal{S})$ be the mapping corresponding to the planning problem such that, for all $(\mu, \gamma) \in \mathcal{S}$,

$$\begin{aligned} \left[\mathcal{T}^W W\right](\mu, \gamma) &= \max_{f^{SP}} \int^{f^{SP}} \left[\frac{1}{a}\left(1 - e^{-a\mu + \frac{a^2}{2\gamma}}\right) - \tilde{f}\right] F\left(d\tilde{f}\right) \\ &\quad + \beta \int \left[R(W)\right]\left(\rho_\theta \mu + s\left(n^{SP}, \gamma\right)\varepsilon, \Gamma\left(n^{SP}, \gamma\right)\right)\Phi(d\varepsilon) \end{aligned}$$

where $n^{SP} = F\left(f^{SP}\right)$.

The following definition introduces the set $\mathcal{W}$ in which the constrained planner's value function lies.

Definition 11. Let $\mathcal{W} \subset B(\mathcal{S})$ be the set of bounded continuous functions $W : (\mu, \gamma) \in \mathcal{S} \longrightarrow \mathbb{R}$ such that W is Lipschitz continuous of modulus $\overline{W}_\mu$ in μ and $\overline{W}_\gamma$ in γ with $\overline{W}_\mu = (1 - \beta\rho_\theta)^{-1} e^{-a\underline{\mu} + \frac{a^2}{2\underline{\gamma}}}$ and $\overline{W}_\gamma = \left(1 - \beta\rho_\theta^2\right)^{-1}\left(\frac{a}{2\underline{\gamma}^2} e^{-a\underline{\mu} + \frac{a^2}{2\underline{\gamma}}} + \beta \|s_\gamma\| \overline{W}_\mu\right)$.

Proposition 3. *Under Assumption 1-3, the recursive competitive equilibrium is constrained inefficient and the efficient allocation can be implemented with investment subsidies $\tau(\mu, \gamma)$ and a uniform tax.*

Proof. Outline: We proceed in two steps. In step 1, after writing down the problem of the social planner and establishing a few regularity conditions such as differentiability (proved in the Online Appendix), we derive the formula that characterizes the planner's optimal cutoff f^{SP}. It differs from the competitive equilibrium cutoff in two ways. First, the planner internalizes the impact of investment on information. Second, because the planner internalizes that investing firms are replaced by new firms, it is not subject to the irreversibility faced by individual firms. We show in step 2 that the planner's cutoff can easily be implemented in the competitive equilibrium with an investment subsidy that captures the social value of information and compensates firms for their option value of waiting.

Step 1. Derivation of the socially optimal cutoff. Given the structure of the problem, it is immediate that the planner's decision takes the form of a cutoff rule f^{SP} such that the planner makes the firms with fixed costs $f \leqslant f^{SP}$ invest. The constrained planner's problem is

$$
\begin{aligned}
W(\mu,\gamma) \quad &= \quad \max_{f^{SP}} \int^{f^{SP}} \left(\mathbb{E}\left[u(\theta) \mid \mu,\gamma\right] - \tilde{f}\right) dF\left(\tilde{f}\right) + \beta \mathbb{E}\left[W\left(\mu',\gamma'\right)\right] \\
& \text{s.t} \quad \mu' = \rho_\theta \frac{\gamma\mu + \gamma_z Z + n^{SP}\gamma_x X}{\gamma + \gamma_z + n^{SP}\gamma_x} \\
& \qquad\; \gamma' = \left(\frac{\rho_\theta^2}{\gamma + \gamma_z + n^{SP}\gamma_x} + \left(1-\rho_\theta^2\right)\sigma_\theta^2\right)^{-1} \\
& \qquad\; n^{SP} = F\left(f^{SP}\right)
\end{aligned} \tag{38}
$$

with $\theta' = \rho_\theta\theta + \varepsilon^\theta, \varepsilon^\theta \sim \mathcal{N}\left(0,\left(1-\rho_\theta^2\right)\sigma_\theta^2\right)$, $Z = \theta + \varepsilon^z, \varepsilon^z \sim \mathcal{N}\left(0,\gamma_z^{-1}\right)$ and $X = \theta + \varepsilon^X, \varepsilon^X \sim \mathcal{N}\left(0,\left(n^{SP}\gamma_x\right)^{-1}\right)$. Using Assumption 2 for the μ process, the value of the planner is the fixed point of the mapping $\mathcal{T}^W$ defined in Definition 10 above. We establish in Lemma 10 in the Online Appendix that this mapping is a well-defined contraction on the space of continuous bounded functions $B(\mathcal{S})$ equipped with the sup norm. Moreover, Lemma 11 shows that $\mathcal{T}^W$ is a self-map on $\mathcal{W}$, which, being a closed subset of a Banach space is itself a Banach space. Hence, the planner's value – the unique fixed point of mapping $\mathcal{T}^W$, must belong to $\mathcal{W}$ and is therefore Lipschitz continuous in μ and γ. Finally, Lemma 12 establishes that the planner's value function is differentiable. The first order condition with respect to the cutoff is

$$
F'\left(f^{SP}\right)\left(\mathbb{E}\left[u(\theta) \mid \mu,\gamma\right] - f^{SP} + \beta \frac{d}{dn^{SP}} \mathbb{E}\left[W\left(\rho_\theta\mu + s\left(n^{SP},\gamma\right)\varepsilon, \Gamma\left(n^{SP},\gamma\right)\right)\right]\right) = 0,
$$

where ε is a standard normal. Since $F' > 0$ under Assumption 1, the efficient cutoff is

$$
f^{SP}(\mu,\gamma) = \mathbb{E}\left[u(\theta) \mid \mu,\gamma\right] + \beta \frac{d}{dn^{SP}} \mathbb{E}\left[W\left(\rho_\theta\mu + s\left(n^{SP},\gamma\right)\varepsilon, \Gamma\left(n^{SP},\gamma\right)\right)\right]. \tag{39}
$$

Step 2. Implementation with subsidies. We show that this optimal cutoff is implementable using belief-dependent investment subsidies $\tau(\mu,\gamma)$ and a uniform tax $T(\mu,\gamma)$ levied on all firms at the beginning of each period to ensure a balanced budget. In the presence of these policy instruments, the firm's problem is

$$
V^\tau(\mu,\gamma,f) = \max\left\{\mathbb{E}\left[u(\theta) \mid \mu,\gamma\right] - f + \tau(\mu,\gamma), \beta \mathbb{E}\left[V^\tau\left(\mu',\gamma',f'\right)\right]\right\} - T(\mu,\gamma), \tag{40}
$$

which yields the individual cutoff rule $f^{c,\tau}$:

$$
f^{c,\tau}(\mu,\gamma) = \mathbb{E}\left[u(\theta) \mid \mu,\gamma\right] + \tau(\mu,\gamma) - \beta \mathbb{E}\left[V^\tau\left(\mu',\gamma',f'\right)\right]. \tag{41}
$$

Requiring that the government's budget constraint balances implies $T(\mu,\gamma) = \tau(\mu,\gamma) F\left(f^{c,\tau}(\mu,\gamma)\right)$.

Comparing (39) and (41), the subsidy implementing the constrained planner's allocation necessarily satisfies

$$\tau(\mu,\gamma) = \beta \frac{d}{dn^{SP}} \mathbb{E}\left[W\left(\rho_\theta \mu + s\left(n^{SP},\gamma\right)\varepsilon, \Gamma\left(n^{SP},\gamma\right)\right)\right] + \beta \mathbb{E}\left[V^\tau\left(\mu',\gamma',f'\right)\right]. \tag{42}$$

This expression is intuitive. The optimal subsidy compensates investors for the two margins that they do not internalize. The first term internalizes the social benefits of adding a marginal signal about θ, which is positive according to Lemma 13. The second term compensates investors for their option value of waiting: since investing firms are replaced by new firms with an investment opportunity, the social planner problem is not subject to the irreversibility that individual investors face. Expression (42) is a functional equation in τ because V^τ depends implicitly on τ. Lemma (14) in the Online Appendix shows that this functional equation admits a unique solution and that the resulting subsidy $\tau(\mu,\gamma)$ and tax $T(\mu,\gamma)$ implement the planner's allocation. □

F.8 Proposition 4

Proposition 4 (Full)**.** *Under Assumptions 1-2, when γ_x is small and if parameters are such that*[49]

$$\left\|F'\right\| \frac{a}{3\overline{\gamma}^2} e^{-a\overline{\mu}+\frac{a^2}{2\overline{\gamma}}} \underline{\Gamma_n} + \underline{\Gamma_\gamma} > 1, \tag{43}$$

then the planner's allocation is subject to uncertainty traps.

Proof. Outline: The proof closely follows Proposition 2. In step 1, we show that there exists a stationary point (μ^*,γ^*), i.e. a point satisfying $\varphi\left(n^{SP}(\mu^*,\gamma^*),\gamma^*\right) = 0$ for φ defined in (37), such that the function $\gamma \mapsto \varphi\left(n^{SP}(\mu^*,\gamma),\gamma\right)$ crosses the x-axis from below. The idea behind the result is identical to Proposition 2: we show that the optimal cutoff f^{SP} increases with γ at some positive rate bounded away from 0 as long as γ_x is small enough that the conditions of Lemma 15 are satisfied. This ensures that we can find a point where $\varphi\left(F\left(f^{SP}(\mu^*,\gamma^*)\right),\gamma^*\right) = 0$ and f^{SP} is the mode of the distribution of fixed costs. If F is sufficiently concentrated around its mode, the function φ crosses the x-axis from below. Step 2 is identical to Proposition 2, so it is omitted.

Step 1. We first assume that γ_x is small enough to satisfy the conditions of Lemma 15, so that f^{SP} increases at least at rate $\underline{f}^{SP}_\mu = \frac{1}{3}e^{-a\overline{\mu}+\frac{a^2}{2\overline{\gamma}}}$ in μ and at rate $\underline{f}^{SP}_\gamma = \frac{a}{3\overline{\gamma}^2}e^{-a\overline{\mu}+\frac{a^2}{2\overline{\gamma}}}$ in γ. We then repeat the same steps as in Proposition 2. Denote $f_m = \operatorname{argmax} F'$. The Theorem of the Maximum tells us that f^{SP} is continuous. From the definition of $\left\{\underline{\gamma},\overline{\gamma}\right\}$ and the continuity of f^{SP}, F and and Γ, there exists $\gamma^* \in \left[\underline{\gamma},\overline{\gamma}\right]$ such that $\varphi\left(F\left(f_m\right),\gamma^*\right) = 0$. Since f^{SP} increases at least at rate $\underline{f}^{SP}_\mu > 0$ and is Lipschitz continuous, we know that there exists $\mu^* \in \left[\underline{\mu},\overline{\mu}\right]$ such that $f^{SP}(\mu^*,\gamma^*) = f_m$ as long as the bounds $\underline{\mu}$ and $\overline{\mu}$ are chosen large enough. By construction, γ^* is a stationary point in the dynamics of γ at $\mu = \mu^*$, i.e., $\varphi\left(n^{SP}(\mu^*,\gamma^*),\gamma^*\right) = 0$. We now show that the function $\gamma \mapsto \varphi\left(n^{SP}(\mu^*,\gamma),\gamma\right)$ crosses the x-axis at γ^* from below. As in Proposition 2, pick $\epsilon > 0$ arbitrarily small; by continuity, there exists a neighborhood Ω of γ^* such that

[49]Condition (43) is satisfied if the dispersion σ_f of the fixed cost distribution is small enough that $\|F'\|$ is large or if the risk aversion parameter a is high enough that $\frac{a}{\overline{\gamma}^2}e^{-a\overline{\mu}+\frac{a^2}{2\overline{\gamma}}}$ is large.

$\inf_{\gamma\in\Omega} F'\left(f^{SP}\left(\mu^*,\gamma\right)\right) \geqslant (1-\varepsilon)\left\|F'\right\|$. Hence, for all $\gamma \in \Omega$,

$$\begin{aligned}\varphi\left(n\left(\mu^*,\gamma\right),\gamma\right) &\geqslant \underline{\varphi}_{\gamma}^{SP,\epsilon}\left(\gamma-\gamma^*\right), \text{ for } \gamma \geqslant \gamma^*,\\ \varphi\left(n\left(\mu^*,\gamma\right),\gamma\right) &\leqslant \underline{\varphi}_{\gamma}^{SP,\epsilon}\left(\gamma-\gamma^*\right), \text{ for } \gamma \leqslant \gamma^*,\end{aligned}$$

where $\underline{\varphi}_{\gamma}^{SP,\varepsilon} = (1-\varepsilon)\left\|F'\right\| \frac{a}{3\overline{\gamma}^2} e^{-a\overline{\mu}+\frac{a^2}{2\overline{\gamma}}} \underline{\Gamma_n} + \underline{\Gamma_{\gamma}} - 1$. Under condition (43) and for ϵ sufficiently small, we have $\underline{\varphi}_{\gamma}^{SP,\epsilon} > 0$. We have thus shown that function $\gamma \mapsto \varphi\left(n\left(\mu^*,\gamma\right),\gamma\right)$ crosses the x-axis at γ^* from below. Step 2 of Proposition 2 applies identically. Hence, we may conclude that there exists a non-empty interval $\left(\mu_l,\mu_h\right)$ such that for all $\mu \in \left(\mu_l,\mu_h\right)$, there exists $\gamma_l\left(\mu\right) < \gamma_h\left(\mu\right)$ such that $\varphi\left(n\left(\mu,\gamma_l\left(\mu\right)\right),\gamma_l\left(\mu\right)\right) = \varphi\left(n\left(\mu,\gamma_h\left(\mu\right)\right),\gamma_h\left(\mu\right)\right) = 0$. Therefore, the social planner's allocation is also subject to uncertainty traps. □

G Auxiliary Lemmas (ONLINE APPENDIX)

This online appendix contains the proofs of the technical lemmas used in the previous proofs.

G.1 Lemma for Belief Dynamics

Lemma 3. *Functions $s\left(n,\gamma\right)$ and $\Gamma\left(n,\gamma\right)$ satisfy:*

1. *The first derivatives of s are such that*

$$\begin{aligned}s_n\left(n,\gamma\right) &= \frac{\rho_\theta\gamma_x}{2\left(\gamma_z+n\gamma_x\right)^{\frac{1}{2}}}\frac{\gamma^{\frac{1}{2}}}{\left(\gamma+\gamma_z+n\gamma_x\right)^{\frac{3}{2}}} \geqslant 0,\\ s_\gamma\left(n,\gamma\right) &= -\frac{\rho_\theta\left(\gamma_z+n\gamma_x\right)^{\frac{1}{2}}}{2}\frac{2\gamma+\gamma_z+n\gamma_x}{\gamma^{\frac{3}{2}}\left(\gamma+\gamma_z+n\gamma_x\right)^{\frac{3}{2}}} \leqslant 0.\end{aligned}$$

   In addition, $s_{nn} \leqslant 0$, $s_{\gamma\gamma} \geqslant 0$, while the sign of $s_{n\gamma}$ varies with n and γ.

2. *The first derivatives of Γ are such that*

$$\begin{aligned}\Gamma_n\left(n,\gamma\right) &= \frac{\rho_\theta^2\gamma_x}{\left(\gamma+\gamma_z+n\gamma_x\right)^2}\Gamma\left(n,\gamma\right)^2 \geqslant 0,\\ \Gamma_\gamma\left(n,\gamma\right) &= \frac{\rho_\theta^2}{\left(\gamma+\gamma_z+n\gamma_x\right)^2}\Gamma\left(n,\gamma\right)^2 \geqslant 0.\end{aligned}$$

   In addition, $\Gamma_{nn} \leqslant 0$, $\Gamma_{\gamma\gamma} \leqslant 0$ and $\Gamma_{n\gamma} \leqslant 0$.

Proof. 1. Compute the first derivatives:

$$\begin{aligned}s_n\left(n,\gamma\right) &= \frac{\rho_\theta\gamma_x}{2}\frac{1}{\left(\gamma+\gamma_z+n\gamma_z\right)^2}\left(\frac{1}{\gamma}-\frac{1}{\gamma+\gamma_z+n\gamma_x}\right)^{-\frac{1}{2}}\\ &= \frac{\rho_\theta\gamma_x}{2\left(\gamma_z+n\gamma_x\right)^{\frac{1}{2}}}\frac{\gamma^{\frac{1}{2}}}{\left(\gamma+\gamma_z+n\gamma_x\right)^{\frac{3}{2}}} \geqslant 0\end{aligned}$$

$$\begin{aligned} s_\gamma(n,\gamma) &= -\frac{\rho_\theta}{2}\left(\frac{1}{\gamma^2}-\frac{1}{(\gamma+\gamma_z+n\gamma_x)^2}\right)\left(\frac{1}{\gamma}-\frac{1}{\gamma+\gamma_z+n\gamma_x}\right)^{-\frac{1}{2}} \\ &= -\frac{\rho_\theta}{2}(\gamma_z+n\gamma_x)^{\frac{1}{2}}\frac{2\gamma+\gamma_z+n\gamma_x}{\gamma^{\frac{3}{2}}(\gamma+\gamma_z+n\gamma_x)^{\frac{3}{2}}}\leqslant 0. \end{aligned}$$

Turning to the second derivatives,

$$s_{nn}(n,\gamma) = -\frac{\rho_\theta\gamma_x^2}{4}\frac{\gamma^{\frac{1}{2}}(\gamma+4\gamma_z+4n\gamma_x)}{(\gamma_z+n\gamma_x)^{\frac{3}{2}}(\gamma+\gamma_z+n\gamma_x)^{\frac{5}{2}}}\leqslant 0$$

$$s_{\gamma\gamma}(n,\gamma) = \frac{\rho_\theta(\gamma_z+n\gamma_x)^{\frac{1}{2}}}{2}\frac{\frac{3}{2}(2\gamma+\gamma_z+n\gamma_x)^2-2\gamma(\gamma+\gamma_z+n\gamma_x)}{\gamma^{\frac{5}{2}}(\gamma+\gamma_z+n\gamma_x)^{\frac{5}{2}}}\geqslant 0$$

The derivative $s_{\gamma n}$ is positive for $\gamma\leqslant\frac{1}{2}(\gamma_z+n\gamma_x)$, negative otherwise:

$$s_{n\gamma}(n,\gamma) = \frac{\rho_\theta\gamma_x}{4(\gamma_z+n\gamma_x)^{\frac{1}{2}}}\gamma^{-\frac{1}{2}}(\gamma+\gamma_z+n\gamma_x)^{-\frac{5}{2}}\frac{\gamma_z+n\gamma_x-2\gamma}{(\gamma+\gamma_z+n\gamma_x)^3}.$$

In particular, since $s_{nn}\leqslant 0$, $|s_n|$ reaches its maximum at $n=0$ and $\gamma=\frac{\gamma_z}{2}$. Similarly, since $s_{\gamma\gamma}\geqslant 0$, $|s_\gamma|$ reaches its maximum at $\gamma=\underline{\gamma}$ and $n=\gamma_x^{-1}\left(2\underline{\gamma}-\gamma_z\right)$.

2. Compute the first derivatives:

$$\Gamma_n(n,\gamma) = \frac{\rho_\theta^2\gamma_x}{(\gamma+\gamma_z+n\gamma_x)^2}\Gamma(n,\gamma)^2\geqslant 0,$$

$$\Gamma_\gamma(n,\gamma) = \frac{\rho_\theta^2}{(\gamma+\gamma_z+n\gamma_x)^2}\Gamma(n,\gamma)^2\geqslant 0.$$

Similarly, the second order derivatives are given by

$$\Gamma_{nn}(n,\gamma) = -2\frac{\rho_\theta^2\gamma_x^2}{(\gamma+\gamma_z+n\gamma_x)^3}\Gamma(n,\gamma)^2\left[1-\frac{\rho_\theta^2}{\gamma+\gamma_z+n\gamma_x}\Gamma(n,\gamma)\right]\leqslant 0,$$

$$\Gamma_{\gamma\gamma}(n,\gamma) = -2\frac{\rho_\theta^2}{(\gamma+\gamma_z+n\gamma_x)^3}\Gamma(n,\gamma)^2\left[1-\frac{\rho_\theta^2}{\gamma+\gamma_z+n\gamma_x}\Gamma(n,\gamma)\right]\leqslant 0,$$

$$\Gamma_{n\gamma}(n,\gamma) = -2\frac{\rho_\theta^2\gamma_x}{(\gamma+\gamma_z+n\gamma_x)^3}\Gamma(n,\gamma)^2\left[1-\frac{\rho_\theta^2}{\gamma+\gamma_z+n\gamma_x}\Gamma(n,\gamma)\right]\leqslant 0.$$

□

G.2 Lemmas for Proposition 1

Lemma 4. *Let $a\in\mathbb{R}$. The following is true:*

1. *For all $(x,y)\in\mathbb{R}^2$, $|max(y,a)-max(x,a)|\leqslant|y-x|$.*

2. *If $g : x \in \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous of modulus $\overline{g}_x$, then $x \mapsto max(a, g(x))$ is also Lipschitz continuous of modulus $\overline{g}_x$.*

3. *If $g : \mathcal{S} \to \mathbb{R}$ is Lipschitz continuous in (μ, γ) of modulus $\left(\overline{g}_\mu, \overline{g}_\gamma\right)$, then $R(g)$ is also Lipschitz continuous of modulus $\left(\overline{g}_\mu, \overline{g}_\gamma\right)$.*

Proof. 1. Let $(x, y) \in \mathbb{R}^2$. Assume WLOG that $x \leqslant y$. Then, either i) $x \leqslant y \leqslant a$, $|\max(y, a) - \max(x, a)| = 0 \leqslant |y - x|$, ii) $x \leqslant a < y$, $|\max(y, a) - \max(x, a)| = |y - a| \leqslant |y - x|$, iii) $a < x \leqslant y$, then$|\max(y, a) - \max(x, a)| = |y - x|$.

2. Let $(x, y) \in \mathbb{R}^2$, then by a trivial application of the previous bullet, $|\max(a, g(y)) - \max(a, g(x))| \leqslant |g(y) - g(x)| \leqslant \overline{g}_x |y - x|$.

3. Let $a \in \mathbb{R}$ and $(\mu, \gamma) \in \mathcal{S}$. From the second bullet, $\forall x \in \mathbb{R}, \max(x, a)$ is Lipschitz of modulus 1 in x. Similarly, $\min(x, a) = -\max(-x, -a)$ is Lipschitz of modulus 1 in x. A composition of Lipschitz continuous functions is a Lipschitz continuous function whose modulus is the product of the individual moduli. Hence, $[R(g)](\mu, \gamma) = g\left(\max\left(\min(\mu, \overline{\mu}), \underline{\mu}\right), \max\left(\min(\gamma, \overline{\gamma}), \underline{\gamma}\right)\right)$ is Lipschitz in (μ, γ) of modulus $\left(\overline{g}_\mu, \overline{g}_\gamma\right)$. □

Lemma 5. *Under the conditions of Proposition 1, the system of equations (31) admits a strictly positive solution.*

Proof. The moduli $\left(\overline{f}^c_\mu, \overline{f}^c_\gamma\right)$ are defined as the solution to the system

$$\begin{cases} \overline{f}^c_\mu = \left|S_\mu\left(\underline{\mu}, \underline{\gamma}\right)\right| + \beta \overline{f}^c_\mu + \beta \left\|F'\right\| \left(\|s_n\| \overline{f}^c_\mu + \overline{\Gamma}_n \overline{f}^c_\gamma\right) \overline{f}^c_\mu, \\ \overline{f}^c_\gamma = \left|S_\gamma\left(\underline{\mu}, \underline{\gamma}\right)\right| + \beta \overline{f}^c_\gamma + \beta \|s_\gamma\| \overline{\Delta_s f} + \beta \left\|F'\right\| \left(\|s_n\| \overline{f}^c_\mu + \overline{\Gamma}_n \overline{f}^c_\gamma\right) \overline{f}^c_\gamma. \end{cases}$$

Multiplying the first row by $\overline{f}^c_\gamma$, the second one by $\overline{f}^c_\mu$ and subtracting both yields

$$\overline{f}^c_\gamma = \frac{\left|S_\gamma\left(\underline{\mu}, \underline{\gamma}\right)\right| + \beta \|s_\gamma\| \overline{\Delta_s f}}{\left|S_\mu\left(\underline{\mu}, \underline{\gamma}\right)\right|} \overline{f}^c_\mu.$$

Substituting the latter expression into the first row of our system provides us with a quadratic equation in $\overline{f}^c_\mu$ only,$\Psi\left(\overline{f}^c_\mu\right) = 0$, where

$$\Psi(f) \equiv \left|S_\mu\left(\underline{\mu}, \underline{\gamma}\right)\right| - (1 - \beta) f + \beta \left\|F'\right\| \left(\|s_n\| + \overline{\Gamma}_n \frac{\left|S_\gamma\left(\underline{\mu}, \underline{\gamma}\right)\right| + \beta \|s_\gamma\| \overline{\Delta_s f}}{\left|S_\mu\left(\underline{\mu}, \underline{\gamma}\right)\right|}\right) f^2.$$

This quadratic equation admits two positive real solutions such that $\Psi\left(\overline{f}^c_\mu\right) = 0$, if and only if

$$(1 - \beta)^2 - 4\beta \left\|F'\right\| \left[\|s_n\| \left|S_\mu\left(\underline{\mu}, \underline{\gamma}\right)\right| + \overline{\Gamma}_n \left(\left|S_\gamma\left(\underline{\mu}, \underline{\gamma}\right)\right| + \beta \|s_\gamma\| \overline{\Delta_s f}\right)\right] \geqslant 0,$$

which corresponds to condition (33). In the case where there exists two distinct solutions $\overline{f}^c_{\mu 1} < \overline{f}^c_{\mu 2}$, we pick the largest one. The corresponding bounds on γ are such that $\overline{f}^c_{\gamma 2} > \overline{f}^c_{\gamma 1}$. □

Lemma 6. *Under the conditions of Proposition 1, the mapping $\mathcal{T}$ maps $\mathcal{F}_0$ onto itself.*

Proof. Let $f^c \in \mathcal{F}_0$. To show that $\mathcal{T}f^c$ belongs to $\mathcal{F}_0$, we must prove that $\mathcal{T}f^c$ is bounded by $\overline{f}^c$ and Lipschitz continuous of moduli $\left(\overline{f}^c_\mu, \overline{f}^c_\gamma\right)$.

1. *$\mathcal{T}f^c$ is bounded by $\overline{f}^c$.* Let $(\mu,\gamma) \in \mathcal{S}$. Under assumptions (1) and 3, we have

$$\begin{aligned}
|\mathcal{T}f^c(\mu,\gamma)| &\leqslant \frac{1}{a}(1-\beta) + \beta\mu_f + \frac{1}{a}e^{-a\underline{\mu}+\frac{a^2}{2}\left(\frac{1}{\underline{\gamma}}+\frac{1}{\gamma_x}\right)}\left(1-\beta e^{a(1-\rho)\underline{\mu}-\frac{a^2}{2}\frac{1-\rho_\theta^2}{\underline{\gamma}}+\frac{a^2}{2}(1-\rho_\theta^2)\sigma_\theta^2}\right) \\
&\quad +\beta\left(\int |f|\,dF + \overline{f}^c\right) \\
&\leqslant \overline{f}^c,
\end{aligned}$$

where the last line results from the definition of $\overline{f}^c$.

2. *$\mathcal{T}f^c$ is Lipschitz in μ of modulus $\overline{f}^c_\mu$.* Let $\mu_1 \leqslant \mu_2 \in \left[\underline{\mu},\overline{\mu}\right], \gamma \in \left[\underline{\gamma},\overline{\gamma}\right]$. Denote $n_i = F\left(f^c(\mu_i,\gamma)\right), i=1,2$. We decompose the differential in μ as follows

$$\mathcal{T}f^c(\mu_2,\gamma) - \mathcal{T}f^c(\mu_1,\gamma) = A + \beta(B+C),$$

where

$$\begin{aligned}
A &= S(\mu_2,\gamma) - S(\mu_1,\gamma) \\
B &= \iint \Big[\max\left(f' - [R(f^c)](\rho_\theta\mu_1 + s(n_2,\gamma)\varepsilon, \Gamma(n_2,\gamma)), 0\right) \\
&\quad -\max\left(f' - [R(f^c)](\rho_\theta\mu_2 + s(n_2,\gamma)\varepsilon, \Gamma(n_2,\gamma)), 0\right)\Big]\Phi(d\varepsilon)\,F\left(df'\right) \\
C &= \iint \max\left(f' - [R(f^c)](\rho_\theta\mu_1 + s(n_1,\gamma)\varepsilon, \Gamma(n_1,\gamma)), 0\right)F\left(df'\right)\Phi(d\varepsilon). \\
&\quad - \iint \max\left(f' - [R(f^c)](\rho_\theta\mu_1 + s(n_2,\gamma)\varepsilon, \Gamma(n_2,\gamma)), 0\right)F\left(df'\right)\Phi(d\varepsilon).
\end{aligned}$$

From the Mean Value Theorem, there exists $\tilde{\mu} \in [\mu_1,\mu_2]$ such that $A = S_\mu(\tilde{\mu},\gamma)(\mu_2-\mu_1)$. Hence,

$$|A| \leqslant \left|S_\mu\left(\underline{\mu},\underline{\gamma}\right)\right||\mu_2-\mu_1|,$$

with $S_\mu(\mu,\gamma) = e^{-a\mu+\frac{a^2}{2\gamma}}\left(1-\beta\rho_\theta e^{a(1-\rho_\theta)\mu-\frac{a^2}{2}\frac{1-\rho_\theta^2}{\gamma}+\frac{a^2}{2}(1-\rho_\theta^2)\sigma_\theta^2}\right)$ and where we have used the result that $|S_\mu|$ is decreasing in μ and γ. Turning to term B, the Lipschitz property of f^c is preserved under the max operator (Lemma 4.2). Hence, $\max\left(f' - f^c(\mu,\gamma), 0\right)$ is Lipschitz continuous in μ of modulus $\overline{f}^c_\mu$. Therefore,

$$|B| \leqslant \overline{f}^c_\mu|\mu_2-\mu_1|.$$

To control for term C, we use the result from Lemma 8, which tells us that

$$|C| \leqslant \overline{\Delta_n f} \gamma_x \left| F\left(f^c\left(\mu_2, \gamma\right)\right) - F\left(f^c\left(\mu_1, \gamma\right)\right) \right| \leqslant \overline{\Delta_n f} \left\| F' \right\| \overline{f}^c_\mu \left| \mu_2 - \mu_1 \right|,$$

where $\overline{\Delta_n f} = \|s_n\| \overline{f}^c_\mu + \overline{\Gamma}_n \overline{f}^c_\gamma$. Hence, $\mathcal{T} f^c$ is Lipschitz in μ of modulus $\overline{f}^c_\mu$ if

$$\left| S_\mu \left(\underline{\mu}, \underline{\gamma}\right) \right| + \beta \left(1 + \overline{\Delta_n f} \left\| F' \right\|\right) \overline{f}^c_\mu \leqslant \overline{f}^c_\mu.$$

This condition is indeed satisfied as it corresponds to the first row of the system (31).

3. *$\mathcal{T} f^c$ is Lipschitz in γ of modulus $\overline{f}^c_\gamma$.* Let $\mu \in \left[\underline{\mu}, \overline{\mu}\right], \gamma_1 \leqslant \gamma_2 \in \left[\underline{\gamma}, \overline{\gamma}\right]$. Denote $n_i = F\left(f^c\left(\mu, \gamma_i\right)\right), i = 1, 2$. We decompose the differential in γ as follows

$$\mathcal{T} f^c\left(\mu, \gamma_2\right) - \mathcal{T} f^c\left(\mu, \gamma_1\right) = A + \beta\left(B + C + D\right)$$

where

$$\begin{aligned}
A &= S\left(\mu, \gamma_2\right) - S\left(\mu, \gamma_1\right) \\
B &= \iint \Big[\max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_2\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right) \\
&\quad - \max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_2\right)\varepsilon, \Gamma\left(n_2, \gamma_2\right)\right), 0\right) \Big] \Phi\left(d\varepsilon\right) F\left(df'\right) \\
C &= \iint \Big[\max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_1\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right) \\
&\quad - \max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_2\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right) \Big] \Phi\left(d\varepsilon\right) F\left(df'\right) \\
D &= \iint \Big[\max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_1, \gamma_1\right)\varepsilon, \Gamma\left(n_1, \gamma_1\right)\right), 0\right) \\
&\quad - \max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_1\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right) \Big] F\left(df'\right) \Phi\left(d\varepsilon\right).
\end{aligned}$$

From the Mean Value Theorem,

$$|A| \leqslant \left| S_\gamma \left(\underline{\mu}, \underline{\gamma}\right) \right| \left| \gamma_2 - \gamma_1 \right|.$$

with $S_\gamma\left(\mu, \gamma\right) = \frac{a}{2\gamma^2} e^{-a\mu + \frac{a^2}{2}\left(\frac{1}{\gamma} + \frac{1}{\gamma_x}\right)} \left(1 - \beta \rho_\theta^2 e^{a(1-\rho_\theta)\mu - \frac{a^2}{2}\frac{1-\rho_\theta^2}{\gamma} + \frac{a^2}{2}(1-\rho_\theta^2)\sigma_\theta^2}\right)$ and we have used the fact that $|S_\gamma|$ is decreasing in μ and γ. From Lemma 4.2, the Lipschitz property is preserved under the max operator. Hence, $\max\left(f' - f^c\left(\mu, \gamma\right), 0\right)$ is Lipschitz continuous in μ of modulus $\overline{f}^c_\gamma$. Therefore,

$$|B| \leqslant \overline{f}^c_\gamma \left| \Gamma\left(n_2, \gamma_2\right) - \Gamma\left(n_2, \gamma_1\right) \right| \leqslant \overline{f}^c_\gamma \overline{\Gamma}_\gamma \left| \gamma_2 - \gamma_1 \right| \leqslant \overline{f}^c_\gamma \left| \gamma_2 - \gamma_1 \right|,$$

where we have used the fact that $\overline{\Gamma}_\gamma = \rho_\theta^2 \left(\frac{\gamma}{\underline{\gamma} + \gamma_z}\right)^2 < 1$. To control for terms C and D, we use Lemma 8, which tell us that

$$|C| \leqslant \overline{\Delta_s f} \left| s\left(n_2,\gamma_2\right) - s\left(n_2,\gamma_1\right) \right| \leqslant \overline{\Delta_s f} \left\| s_\gamma \right\| \left| \gamma_2 - \gamma_1 \right|,$$

$$|D| \leqslant \overline{\Delta_n f} \left| F\left(f^c\left(\mu,\gamma_2\right)\right) - F\left(f^c\left(\mu,\gamma_1\right)\right) \right| \leqslant \overline{\Delta_n f} \left\| F' \right\| \overline{f}^c_\gamma \left| \gamma_2 - \gamma_1 \right|,$$

where $\overline{\Delta_s f} = \overline{f^c} \left(\frac{1}{s(0,\overline{\gamma})} + \frac{1}{s(0,\overline{\gamma})^3} \right)$, and $\overline{\Delta_n f} = \|s_n\| \overline{f}^c_\mu + \overline{\Gamma}_n \overline{f}^c_\gamma$. Thus, $\mathcal{T} f^c$ is Lipschitz continuous in γ of modulus $\overline{f}^c_\gamma$ if

$$\left| S_\gamma\left(\underline{\mu},\underline{\gamma}\right) \right| + \beta \overline{\Delta_s f} \left\| s_\gamma \right\| + \beta \overline{f}^c_\gamma + \beta \overline{\Delta_n f} \left\| F' \right\| \overline{f}^c_\gamma \leqslant \overline{f}^c_\gamma,$$

which is satisfied as it corresponds to the second row of the system (31).

□

Lemma 7. *Under the conditions of Proposition 1 and, in particular, conditions (34) and (35), the mapping $\mathcal{T}$ maps $\mathcal{F}_1$ onto itself.*[50] *In addition, the rates of increase of $\mathcal{T} f^c$ are bounded from below as follows:*

$$\forall \quad \mu_2 \geqslant \mu_1, \mathcal{T} f^c\left(\mu_2,\gamma\right) - \mathcal{T} f^c\left(\mu_1,\gamma\right) \geqslant \left(S_\mu\left(\mu_2,\gamma\right) - \beta \left\| F' \right\| \overline{\Delta_n f f}^c_\mu \right) \left(\mu_2 - \mu_1\right),$$
$$\forall \quad \gamma_2 \geqslant \gamma_1, \mathcal{T} f^c\left(\mu,\gamma_2\right) - \mathcal{T} f^c\left(\mu,\gamma_1\right) \geqslant \left(S_\gamma\left(\mu,\gamma_2\right) - \beta \left\| s_\gamma \right\| \overline{\Delta_s f} - \beta \left\| F' \right\| \overline{\Delta_n f f}^c_\mu \right) \left(\gamma_2 - \gamma_1\right).$$

Proof. Let $f^c \in \mathcal{F}_1$. Denote $f^c_n\left(\mu,\gamma\right) \equiv \iint \max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n,\gamma\right)\varepsilon, \Gamma\left(n,\gamma\right)\right), 0\right)\Big] \Phi\left(d\varepsilon\right) F\left(df'\right)$. We prove each property step by step.

1. $\mathcal{T} f^c$ *is weakly increasing in* μ. Let $\left(\mu_1,\mu_2\right) \in \left[\underline{\mu},\overline{\mu}\right]^2$ with $\mu_1 < \mu_2$ and $\gamma \in \left[\underline{\gamma},\overline{\gamma}\right]$. Denote $n_i = F\left(f^c\left(\mu_i,\gamma\right)\right), i = 1,2$. Using the same decomposition as in Lemma 6.2:

$$\left[\mathcal{T} f^c\right]\left(\mu_2,\gamma\right) - \left[\mathcal{T} f^c\right]\left(\mu_1,\gamma\right) = A + \beta\left(B + C\right)$$

where

$$\begin{aligned} A &= S\left(\mu_2,\gamma\right) - S\left(\mu_1,\gamma\right) \\ B &= f^c_{n_2}\left(\mu_1,\gamma\right) - f^c_{n_2}\left(\mu_2,\gamma\right) \\ C &= f^c_{n_1}\left(\mu_1,\gamma\right) - f^c_{n_2}\left(\mu_1,\gamma\right). \end{aligned}$$

Since $S_\mu > 0$ and $S_{\mu\mu} \leqslant 0$, we have that

$$A \geqslant S_\mu\left(\mu_2,\gamma\right)\left(\mu_2 - \mu_1\right) \geqslant S_\mu\left(\overline{\mu},\overline{\gamma}\right)\left(\mu_2 - \mu_1\right).$$

[50] Conditions (34) and (35) are satisfied when γ_x is small enough that $\overline{\Delta_n f}$ is negligible and if the degree of risk aversion a is large enough that $S_\gamma > \beta \overline{\Delta_s f} \left\| s_\gamma \right\|$.

Turning to term B, since $f^c \in \mathcal{F}_1$ is increasing in μ, we have $B \geqslant 0$. Term C can be controlled in the same way as in Lemma 6.2:

$$|C| \leqslant \overline{\Delta_n f} \left\|F'\right\| \overline{f}^c_\mu \left(\mu_2 - \mu_1\right).$$

Hence, $\mathcal{T} f^c$ increases at least at rate

$$\forall \mu_2 \geqslant \mu_1 \in \left[\underline{\mu}, \overline{\mu}\right], \mathcal{T} f^c \left(\mu_2, \gamma\right) - \mathcal{T} f^c \left(\mu_1, \gamma\right) \geqslant \left(S_\mu \left(\mu_2, \gamma\right) - \beta \overline{\Delta_n f} \left\|F'\right\| \overline{f}^c_\mu\right) \left(\mu_2 - \mu_1\right),$$

and $\mathcal{T} f^c$ is weakly increasing in μ as long as

$$S_\mu \left(\overline{\mu}, \overline{\gamma}\right) \geqslant \beta \overline{\Delta_n f} \left\|F'\right\| \overline{f}^c_\mu,$$

which is satisfied for γ_x low such that $\overline{\Delta_n f}$ is small enough.

2. $\mathcal{T} f^c$ *is weakly increasing in* γ. Let $\mu \in \left[\underline{\mu}, \overline{\mu}\right]$ and $\gamma_1 \leqslant \gamma_2 \in \left[\underline{\gamma}, \overline{\gamma}\right]$. Denote $n_i = F\left(f^c\left(\mu, \gamma_i\right)\right), i = 1, 2$. Using the same decomposition as in Lemma 6.3:

$$\mathcal{T} f^c \left(\mu, \gamma_2\right) - \mathcal{T} f^c \left(\mu, \gamma_1\right) = A + \beta \left(B + C + D\right)$$

where

$$\begin{aligned}
A &= S\left(\mu, \gamma_2\right) - S\left(\mu, \gamma_1\right) \\
B &= \iint \Big[\max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_2\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right) \\
&\quad - \max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_2\right)\varepsilon, \Gamma\left(n_2, \gamma_2\right)\right), 0\right)\Big] \Phi\left(d\varepsilon\right) F\left(df'\right) \\
C &= \iint \Big[\max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_1\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right) \\
&\quad - \max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma_2\right)\varepsilon, \Gamma\left(n_2, \gamma_1\right)\right), 0\right)\Big] \Phi\left(d\varepsilon\right) F\left(df'\right) \\
D &= f^c_{n_1}\left(\mu_1, \gamma_1\right) - f^c_{n_2}\left(\mu, \gamma_1\right).
\end{aligned}$$

Since $S_\gamma > 0$ and $S_{\gamma\gamma} \leqslant 0$, we have that

$$A \geqslant S_\gamma \left(\overline{\mu}, \overline{\gamma}\right) \left(\mu_2 - \mu_1\right) > 0.$$

Turning to term B, we use the fact that $f^c \in \mathcal{F}_1$ is weakly increasing in γ. Since $\Gamma\left(n_2, \gamma_1\right) \leqslant \Gamma\left(n_2, \gamma_2\right)$, we have that $B \geqslant 0$. We control terms C and D in the same way we did in Lemma 6.3:

$$|C| \quad \leqslant \quad \overline{\Delta_s f} \left\|s_\gamma\right\| \left|\gamma_2 - \gamma_1\right|$$

$$|D| \leqslant \overline{\Delta_n f} \, \|F'\| \, \overline{f}^c_\gamma \left(\gamma_2 - \gamma_1\right),$$

where $\overline{\Delta_s f} = \overline{f^c}\left(\frac{1}{s(0,\overline{\gamma})} + \frac{1}{s(0,\overline{\gamma})^3}\right)$, and $\overline{\Delta_n f} = \|s_n\| \, \overline{f}^c_\mu + \overline{\Gamma}_n \overline{f}^c_\gamma$. Hence, $\mathcal{T} f^c$ increases in γ at least at rate

$$\forall \gamma_2 \geqslant \gamma_1 \in \left[\underline{\gamma}, \overline{\gamma}\right], \mathcal{T} f^c\left(\mu, \gamma_2\right) - \mathcal{T} f^c\left(\mu, \gamma_1\right) \geqslant \left(S_\gamma\left(\mu, \gamma_2\right) - \beta \left\|s_\gamma\right\| \overline{\Delta_s f} - \beta \left\|F'\right\| \overline{\Delta_n f} \overline{f}^c_\mu\right)\left(\gamma_2 - \gamma_1\right),$$

and $\mathcal{T} f^c$ is weakly increasing in γ as long as

$$S_\gamma\left(\overline{\mu}, \overline{\gamma}\right) \geqslant \beta\left(\overline{\Delta_s f} \left\|s_\gamma\right\| + \overline{\Delta_n f} \left\|F'\right\| \overline{f}^c_\gamma\right),$$

which corresponds to condition 34 and is satisfied whenever risk aversion is large enough that $S_\gamma\left(\overline{\mu}, \overline{\gamma}\right) > \beta \overline{\Delta_s f} \left\|s_\gamma\right\|$ and γ_x is small enough that $\overline{\Delta_n f}$ is negligible.

□

Lemma 8. *For $f^c \in \mathcal{F}_0$, denote $f^c_n\left(\mu, \gamma\right) \equiv \iint max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n, \gamma\right) \varepsilon, \Gamma\left(n, \gamma\right)\right), 0\right) F\left(df'\right) \Phi\left(d\varepsilon\right)$. The following inequalities hold.*

1. *For $n_1, n_2 \geqslant 0$, $\left\|f^c_{n_2} - f^c_{n_1}\right\| \leqslant \overline{\Delta_n f} \left|n_2 - n_1\right|$ where $\overline{\Delta_n f} = \|s_n\| \, \overline{f}^c_\mu + \overline{\Gamma}_n \overline{f}^c_\gamma$;*
2. *For $s_1, s_2 \geqslant 0$,*

$$\left|\iint\left[max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s_2 \varepsilon, \Gamma\left(n, \gamma\right)\right), 0\right) - max\left(f' - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s_1 \varepsilon, \Gamma\left(n, \gamma\right)\right), 0\right)\right] F\left(df'\right) \Phi\left(d\varepsilon\right)\right| \leqslant \overline{\Delta_s f} \left|s_2 - s_1\right|,$$

with $\overline{\Delta_s f} = \overline{f}^c\left(\frac{1}{s(0,\overline{\gamma})} + \frac{1}{s(0,\overline{\gamma})^3}\right)$.

Proof. Fix $\mu \in \left[\underline{\mu}, \overline{\mu}\right]$ and $\gamma \in \left[\underline{\gamma}, \overline{\gamma}\right]$.

1. Since the Lipschitz property is preserved under the max and R operators (Lemma 4), $\max\left(f' - \left[R\left(f^c\right)\right]\left(\mu', \gamma'\right), 0\right)$ is Lipschitz continuous of moduli $\left(\overline{f}^c_\mu, \overline{f}^c_\gamma\right)$ in μ and γ. Hence,

$$\begin{aligned}
& \left|f^c_{n_2}\left(\mu, \gamma\right) - f^c_{n_1}\left(\mu, \gamma\right)\right| \\
\leqslant & \int \left|\left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_2, \gamma\right) \varepsilon, \Gamma\left(n_2, \gamma\right)\right) - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_1, \gamma\right) \varepsilon, \Gamma\left(n_2, \gamma\right)\right)\right| \Phi\left(d\varepsilon\right) \\
& + \int \left|\left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_1, \gamma\right) \varepsilon, \Gamma\left(n_2, \gamma\right)\right) - \left[R\left(f^c\right)\right]\left(\rho_\theta \mu + s\left(n_1, \gamma\right) \varepsilon, \Gamma\left(n_1, \gamma\right)\right)\right| \Phi\left(d\varepsilon\right) \\
\leqslant & \overline{f}^c_\mu \left|s\left(n_2, \gamma\right) - s\left(n_1, \gamma\right)\right| \int \left|\varepsilon\right| \Phi\left(d\varepsilon\right) + \overline{f}^c_\gamma \left|\Gamma\left(n_2, \gamma\right) - \Gamma\left(n_1, \gamma\right)\right|. \\
\leqslant & \left(\overline{f}^c_\mu \left\|s_n\right\| + \overline{f}^c_\gamma \overline{\Gamma}_n\right)\left|n_2 - n_1\right|,
\end{aligned}$$

where we have used $\int \left|\varepsilon\right| \Phi\left(d\varepsilon\right) = \sqrt{\frac{2}{\pi}} < 1$.

2. Assume WLOG that $s_1 \leqslant s_2$. Using Lemma 4, the difference between the maximums is bounded by

$$\begin{aligned} A &\equiv \left| \iint \left[\max \left(f' - \left[R\left(f^c \right) \right] \left(\rho_\theta \mu + s_2 \varepsilon, \Gamma\left(n, \gamma \right) \right), 0 \right) \right. \right. \\ &\quad \left. \left. - \max \left(f' - \left[R\left(f^c \right) \right] \left(\rho_\theta \mu + s_1 \varepsilon, \Gamma\left(n, \gamma \right) \right), 0 \right) \right] F\left(df' \right) \Phi\left(d\varepsilon \right) \right| \\ &\leqslant \left| \iint \left[\left[R\left(f^c \right) \right] \left(\rho_\theta \mu + s_2 \varepsilon, \Gamma\left(n, \gamma \right) \right) - \left[R\left(f^c \right) \right] \left(\rho_\theta \mu + s_1 \varepsilon, \Gamma\left(n, \gamma \right) \right) \right] F\left(df' \right) \Phi\left(d\varepsilon \right) \right|. \end{aligned}$$

We then proceed with the change of variable $\mu' = \rho_\theta \mu + s\left(n, \gamma\right) \varepsilon$,

$$\iint \left[R\left(f^c \right) \right] \left(\rho_\theta \mu + s_i \varepsilon, \Gamma\left(n, \gamma \right) \right) F\left(df' \right) \Phi\left(d\varepsilon \right) = \iint \left[R\left(f^c \right) \right] \left(\mu', \Gamma\left(n, \gamma \right) \right) \frac{1}{s_i \sqrt{2\pi}} e^{-\frac{\left(\mu' - \rho_\theta \mu \right)^2}{2 s_i^2}} F\left(df' \right) d\mu'.$$

Hence,

$$A \leqslant \left| \iint \left[R\left(f^c \right) \right] \left(\mu', \Gamma\left(n, \gamma \right) \right) \left(\frac{1}{s_2 \sqrt{2\pi}} e^{-\frac{\left(\mu' - \rho_\theta \mu \right)^2}{2 s_i^2}} - \frac{1}{s_1 \sqrt{2\pi}} e^{-\frac{\left(\mu' - \rho_\theta \mu \right)^2}{2 s_i^2}} \right) F\left(df' \right) d\mu' \right|.$$

From the fundamental theorem of calculus and summing over s,

$$\begin{aligned} A &\leqslant \left| \iiint_{s_1}^{s_2} \left[R\left(f^c \right) \right] \left(\mu', \Gamma\left(n, \gamma \right) \right) \left(-\frac{1}{s^2} + \frac{\left(\mu' - \rho_\theta \mu \right)^2}{s^4} \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(\mu' - \rho_\theta \mu \right)^2}{2 s^2}} F\left(df' \right) d\mu' ds \right| \\ &\leqslant \left| \iiint_{s_1}^{s_2} \left[R\left(f^c \right) \right] \left(\rho_\theta \mu + s\varepsilon, \Gamma\left(n, \gamma \right) \right) \left(-\frac{1}{s} + \frac{\varepsilon^2}{s^3} \right) F\left(df' \right) \Phi\left(d\varepsilon \right) ds \right| \\ &\leqslant \overline{f}^c \iiint_{s_1}^{s_2} \left| -\frac{1}{s} + \frac{\varepsilon^2}{s^3} \right| F\left(df' \right) \Phi\left(d\varepsilon \right) ds \\ &\leqslant \overline{f}^c \left(\frac{1}{s\left(0, \overline{\gamma} \right)} + \frac{1}{s\left(0, \overline{\gamma} \right)^3} \right) \left| s_2 - s_1 \right|. \end{aligned}$$

□

G.3 Lemmas for Proposition 2

Lemma 9. *Under the conditions of Lemma 7, the equilibrium cutoff $f^c \in \mathcal{F}_1$ is strictly increasing in μ and γ and satisfies $\forall \mu_1 \leqslant \mu_2 \in \left[\underline{\mu}, \overline{\mu} \right], \gamma_1 \leqslant \gamma_2 \in \left[\underline{\gamma}, \overline{\gamma} \right]$:*

$$\begin{aligned} F\left(f^c\left(\mu_2, \gamma_1 \right) \right) - F\left(f^c\left(\mu_1, \gamma_1 \right) \right) &\geqslant \left(1 + \beta \overline{f}^c_\mu \left\| F' \right\| \left\| s_n \right\| \right)^{-1} S_\mu\left(\overline{\mu}, \overline{\gamma} \right) \left[\min_{\left[f^c\left(\mu_1, \gamma_1 \right), f^c\left(\mu_2, \gamma_1 \right) \right]} F' \right] \left(\mu_2 - \mu_1 \right), \\ F\left(f^c\left(\mu_1, \gamma_2 \right) \right) - F\left(f^c\left(\mu_1, \gamma_1 \right) \right) &\geqslant \left(1 + \beta \overline{f}^c_\mu \left\| F' \right\| \left\| s_n \right\| \right)^{-1} \left[S_\gamma\left(\overline{\mu}, \overline{\gamma} \right) - \beta \overline{\Delta_s f} \left\| s_\gamma \right\| \right] \left[\min_{\left[f^c\left(\mu_1, \gamma_1 \right), f^c\left(\mu_1, \gamma_2 \right) \right]} F' \right] \left(\gamma_2 - \gamma_1 \right). \end{aligned}$$

Proof. Let us define the following two auxiliary functions ψ and Ψ as

$$\begin{cases} \psi\left(\mu,\gamma,\tilde{n}\right) & = S\left(\mu,\gamma\right)-\beta\iint\max\left[f'-R\left(f^{c}\right)\left(\rho_{\theta}\mu+s\left(\tilde{n},\gamma\right)\varepsilon,\Gamma\left(F\left(f^{c}\left(\mu,\gamma\right)\right),\gamma\right)\right),0\right]F\left(df'\right)\Phi\left(d\varepsilon\right), \\ \Psi\left(\mu,\gamma,\tilde{n}\right) & = F\left(\psi\left(\mu,\gamma,\tilde{n}\right)\right)-\tilde{n}. \end{cases}$$

By construction, since f^c is the equilibrium cutoff, we have

$$\forall\left(\mu,\gamma\right)\in\mathcal{S},\Psi\left(\mu,\gamma,F\left(f^{c}\left(\mu,\gamma\right)\right)\right)=0,$$

and similarly, $\psi\left(\mu,\gamma,F\left(f^{c}\left(\mu,\gamma\right)\right)\right)=f^{c}\left(\mu,\gamma\right),\forall\left(\mu,\gamma\right)\in\mathcal{S}$. Note that, from the properties of $f^c\in\mathcal{F}^1$, Ψ is Lipschitz continuous in μ and γ. In addition, since $S_\mu>0$, $S_\gamma>0$, $s_\gamma<0$, $\Gamma_n>0$, $\Gamma_\gamma>0$ and $f^c\in\mathcal{F}_1$ is monotonic, Ψ is strictly increasing in μ and γ such that

$$\begin{aligned} &\forall\mu_{1}\leqslant\mu_{2}\in\left[\underline{\mu},\overline{\mu}\right],\gamma\in\left[\underline{\gamma},\overline{\gamma}\right], \\ &\qquad \Psi\left(\mu_{2},\gamma,\tilde{n}\right)-\Psi\left(\mu_{1},\gamma,\tilde{n}\right) \geqslant S_{\mu}\left(\overline{\mu},\overline{\gamma}\right)\left[\underset{\left[\psi\left(\mu_{1},\gamma,\tilde{n}\right),\psi\left(\mu_{2},\gamma,\tilde{n}\right)\right]}{\min F'}\right]\left(\mu_{2}-\mu_{1}\right), \\ &\forall\mu\in\left[\underline{\mu},\overline{\mu}\right],\gamma_{1}\leqslant\gamma_{2}\in\left[\underline{\gamma},\overline{\gamma}\right], \\ &\qquad \Psi\left(\mu,\gamma_{2},\tilde{n}\right)-\Psi\left(\mu,\gamma_{1},\tilde{n}\right) \geqslant \left[S_{\gamma}\left(\overline{\mu},\overline{\gamma}\right)-\beta\overline{\Delta_{s}f}\left\Vert s_{\gamma}\right\Vert\right]\left[\underset{\left[\psi\left(\mu,\gamma_{1},\tilde{n}\right),\psi\left(\mu,\gamma_{2},\tilde{n}\right)\right]}{\min F'}\right]\left(\gamma_{2}-\gamma_{1}\right). \end{aligned}$$

Note that we have used the result that a lower bound on the rate of increase of Ψ as a function of γ is $S_\gamma\left(\overline{\mu},\overline{\gamma}\right)-\beta\overline{\Delta_s f}\left\Vert s_\gamma\right\Vert>0$ according to Lemma 7. On the other hand, using similar arguments to Lemma (8), we have that ψ is Lipschitz in $\tilde{n}$ of modulus $\beta\overline{f}^c_\mu\left\Vert s_n\right\Vert$. Hence, for $\tilde{n}_1\leqslant\tilde{n}_2$, that

$$-\left(1+\beta\overline{f}_{\mu}^{c}\left\Vert F'\right\Vert\left\Vert s_{n}\right\Vert\right)\left(\tilde{n}_{2}-\tilde{n}_{1}\right)\leqslant\Psi\left(\mu,\gamma,\tilde{n}_{2}\right)-\Psi\left(\mu,\gamma,\tilde{n}_{1}\right)\leqslant-\left(1-\beta\overline{f}_{\mu}^{c}\left\Vert F'\right\Vert\left\Vert s_{n}\right\Vert\right)\left(\tilde{n}_{2}-\tilde{n}_{1}\right)\leqslant0,$$

where we have used the fact $\beta\left\Vert F'\right\Vert\left\Vert s_n\right\Vert\overline{f}^c_\mu\leqslant\beta\left\Vert F'\right\Vert\overline{\Delta_n f}<1$ under condition (33), as argued in Proposition 1. The implicit function theorem for Lipschitz functions (see Dontchev and Rockafellar (2009)) tells us that the solution $\tilde{n}\left(\mu,\gamma\right)=F\left(f^c\left(\mu,\gamma\right)\right)$ to equation $\Psi\left(\mu,\gamma,\tilde{n}\left(\mu,\gamma\right)\right)=0$ is such that, $\forall\left(\mu_i,\gamma_i\right)\in\mathcal{S},i=1,2$:

$$\begin{aligned} \tilde{n}\left(\mu_{2},\gamma_{1}\right)-\tilde{n}\left(\mu_{1},\gamma_{1}\right) &\geqslant \left(1+\beta\overline{f}_{\mu}^{c}\left\Vert F'\right\Vert\left\Vert s_{n}\right\Vert\right)^{-1}\left[\underset{\left[f^{c}\left(\mu_{1},\gamma_{1}\right),f^{c}\left(\mu_{2},\gamma_{1}\right)\right]}{\min F'}\right]S_{\mu}\left(\overline{\mu},\overline{\gamma}\right)\left(\mu_{2}-\mu_{1}\right), \\ \tilde{n}\left(\mu_{1},\gamma_{2}\right)-\tilde{n}\left(\mu_{1},\gamma_{1}\right) &\geqslant \left(1+\beta\overline{f}_{\mu}^{c}\left\Vert F'\right\Vert\left\Vert s_{n}\right\Vert\right)^{-1}\left[\underset{\left[f^{c}\left(\mu_{1},\gamma_{1}\right),f^{c}\left(\mu_{1},\gamma_{2}\right)\right]}{\min F'}\right]\left[S_{\gamma}\left(\overline{\mu},\overline{\gamma}\right)-\beta\overline{\Delta_{s}f}\left\Vert s_{\gamma}\right\Vert\right]\left(\gamma_{2}-\gamma_{1}\right). \end{aligned}$$

□

G.4 Lemmas for Proposition 3

Lemma 10. *Under the conditions of Proposition 3, the mapping $\mathcal{T}^W$ is a contraction on $B\left(\mathcal{S}\right)$.*

Proof. Let $W\in B\left(\mathcal{S}\right)$. Note first that $\mathcal{T}^W$ is a well-defined self-map on $B\left(\mathcal{S}\right)$. Indeed, the Theorem of the Maximum tells us that $\mathcal{T}^W W$ is continuous. It is also bounded, since for all $(\mu,\gamma)\in\mathcal{S}$,

$$\left|\left[\mathcal{T}^W W\right](\mu,\gamma)\right| \leqslant \frac{1}{a}\left[1+e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}\right]+\int |f|\, dF+\beta\,\|W\|\,.$$

We then conclude by noticing that mapping $\mathcal{T}^W$ trivially satisfies Blackwell's sufficient conditions. Indeed, $\mathcal{T}^W$ satisfies

1. Monotonicity: Let $W_1 \leqslant W_2 \in B(\mathcal{S})$, then $\mathcal{T}^W W_2 \geqslant \mathcal{T}^W W_1$.
2. Discounting: Let $W\in B(\mathcal{S})$ and $a\in\mathbb{R}$, then $\mathcal{T}^W(W+a)=\mathcal{T}^W W+\beta a$.

Hence, $\mathcal{T}^W$ is a contraction on $B(\mathcal{S})$.

□

Lemma 11. *Under the conditions of Proposition 3, the mapping $\mathcal{T}^W$ is a well-defined self-map on $\mathcal{W}\subset B(\mathcal{S})$.*

Proof. Let $W\in\mathcal{W}$. We must verify that $\mathcal{T}^W W$ is Lipschitz continuous of moduli $\left(\overline{W}_\mu,\overline{W}_\gamma\right)$. In what follows, it is useful to define the value of planning problem when evaluated at some arbitrary cutoff in the current period. Define $\tilde{\mathcal{T}}^W$ such that

$$\begin{aligned}\tilde{\mathcal{T}}^W(W)(\mu,\gamma,f) \;\equiv\; & \int^f\left[\frac{1}{a}\left(1-e^{-a\mu+\frac{a^2}{2\gamma}}\right)-\tilde{f}\right]F\left(d\tilde{f}\right)\\ & +\beta\int\left[R(W)\right]\left(\rho_\theta\mu+s\left(F(f),\gamma\right)\varepsilon,\Gamma\left(F(f),\gamma\right)\right)\Phi(d\varepsilon)\,.\end{aligned}$$

Then, by definition, $\left[\mathcal{T}^W(W)\right](\mu,\gamma)=\max_f\left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma,f)$.

1. *$\mathcal{T}^W W$ is Lipschitz continuous in μ with modulus $\overline{W}_\mu$.* Let $\mu_1\leqslant\mu_2\in\left[\underline{\mu},\overline{\mu}\right]$ and $\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]$. Denote $f_i^{SP}, i=1,2$ the optimal cutoffs and entry associated to the maximization problem when evaluated at μ_i. Let $n_i=F\left(f_i^{SP}\right)$. The difference $\left[\mathcal{T}^W W\right](\mu_2,\gamma)-\left[\mathcal{T}^W W\right](\mu_1,\gamma)$ can be bounded above by evaluating the value of the planning problem for μ_1 at the suboptimal point f_2^{SP}. Hence,

$$\begin{aligned}\left[\mathcal{T}^W W\right](\mu_2,\gamma)-\left[\mathcal{T}^W W\right](\mu_1,\gamma) \;&\leqslant\; \left[\tilde{\mathcal{T}}^W W\right]\left(\mu_2,\gamma,f_2^{SP}\right)-\left[\tilde{\mathcal{T}}^W W\right]\left(\mu_1,\gamma,f_2^{SP}\right)\\ &\leqslant\; A+\beta B\end{aligned}$$

where

$$\begin{aligned}A \;&=\; \left|\int_{\underline{f}}^{{f_2}^{SP}}\left(\mathbb{E}\left[u(\theta)\mid\mu_2,\gamma\right]-\tilde{f}\right)dF\left(\tilde{f}\right)-\int_{\underline{f}}^{{f_2}^{SP}}\left(\mathbb{E}\left[u\left(\theta+\varepsilon^z\right)\mid\mu_1,\gamma\right]-\tilde{f}\right)dF\left(\tilde{f}\right)\right|\\ B \;&=\; \left|\int\left[R(W)\right]\left(\rho_\theta\mu_2+s(n_2,\gamma)\varepsilon,\Gamma(n_2,\gamma)\right)\Phi(d\varepsilon)-\int\left[R(W)\right]\left(\rho_\theta\mu_1+s(n_2,\gamma)\varepsilon,\Gamma(n_2,\gamma)\right)\Phi(d\varepsilon)\right|.\end{aligned}$$

Since $\mathbb{E}\left[u\left(\theta\right)\mid\mu,\gamma\right]=\frac{1}{a}\left[1-e^{-a\mu+\frac{a^2}{2\gamma}}\right]$, we control term A as follows

$$\left|A\right|\leqslant e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}\left|\mu_2-\mu_1\right|.$$

Using Lemma 4.3 and the Lipschitz property of $R\left(W\right)$, we bound B by

$$\left|B\right|\leqslant\rho_\theta\overline{W}_\mu\left|\mu_2-\mu_1\right|.$$

Repeating the same argument for $\left[\mathcal{T}^W W\right]\left(\mu_1,\gamma\right)-\left[\mathcal{T}^W W\right]\left(\mu_2,\gamma\right)$ to obtain the same bounds, we conclude from the definition

$$\overline{W}_\mu=\left(1-\beta\rho_\theta\right)^{-1}e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}$$

that Lipschitz continuity of modulus $\overline{W}_\mu$ is preserved by the mapping $\mathcal{T}^W$.

2. *$\mathcal{T}^W W$ is Lipschitz continuous in γ with modulus $\overline{W}_\gamma$.* Let $\mu\in\left[\underline{\mu},\overline{\mu}\right]$ and $\gamma_1\leqslant\gamma_2\in\left[\underline{\gamma},\overline{\gamma}\right]$. Denote $f_i^{SP},i=1,2$ the optimal cutoffs and entry associated to the maximization problem when evaluated at γ_i. Let $n_i=F\left(f_i^{SP}\right)$. The difference $\left[\mathcal{T}^W W\right]\left(\mu,\gamma_2\right)-\left[\mathcal{T}^W W\right]\left(\mu,\gamma_1\right)$ can be bounded above by evaluating the value of the planning problem for γ_1 at the suboptimal point f_2^{SP}. Hence,

$$\begin{aligned}\left[\mathcal{T}^W W\right]\left(\mu,\gamma_2\right)-\left[\mathcal{T}^W W\right]\left(\mu,\gamma_1\right) &\leqslant \left[\tilde{\mathcal{T}}^W W\right]\left(\mu,\gamma_2,f_2^{SP}\right)-\left[\tilde{\mathcal{T}}^W W\right]\left(\mu,\gamma_1,f_2^{SP}\right)\\ &\leqslant A+\beta\left(B+C\right)\end{aligned}$$

where

$$\begin{aligned}A &= \left|\int_{\underline{f}}^{f_2{}^{SP}}\left(\mathbb{E}\left[u\left(\theta\right)\mid\mu,\gamma_2\right]-\tilde{f}\right)dF\left(\tilde{f}\right)-\int_{\underline{f}}^{f_2{}^{SP}}\left(\mathbb{E}\left[u\left(\theta+\varepsilon^z\right)\mid\mu,\gamma_1\right]-\tilde{f}\right)dF\left(\tilde{f}\right)\right|\\ B &= \left|\int\left[R\left(W\right)\right]\left(\rho_\theta\mu+s\left(n_2,\gamma_2\right)\varepsilon,\Gamma\left(n_2,\gamma_2\right)\right)\Phi\left(d\varepsilon\right)-\int\left[R\left(W\right)\right]\left(\rho_\theta\mu+s\left(n_2,\gamma_1\right)\varepsilon,\Gamma\left(n_2,\gamma_2\right)\right)\Phi\left(d\varepsilon\right)\right|\\ C &= \left|\int\left[R\left(W\right)\right]\left(\rho_\theta\mu+s\left(n_2,\gamma_1\right)\varepsilon,\Gamma\left(n_2,\gamma_2\right)\right)\Phi\left(d\varepsilon\right)-\int\left[R\left(W\right)\right]\left(\rho_\theta\mu+s\left(n_2,\gamma_1\right)\varepsilon,\Gamma\left(n_2,\gamma_1\right)\right)\Phi\left(d\varepsilon\right)\right|.\end{aligned}$$

Using the same arguments as in the previous step, we have

$$\left|A\right|\leqslant\frac{a}{2\underline{\gamma}^2}e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}}\left|\gamma_2-\gamma_1\right|.$$

Using Lemma 4.3, the Lipschitz property of $R\left(W\right)$ and the Mean Value Theorem on s and Γ, we can bound B and C by

$$\begin{aligned}\left|B\right| &\leqslant \overline{W}_\mu\left\|s_\gamma\right\|\int\left|\varepsilon\right|d\Phi\left|\gamma_2-\gamma_1\right|\leqslant\overline{W}_\mu\left\|s_\gamma\right\|\left|\gamma_2-\gamma_1\right|\\ \left|C\right| &\leqslant \overline{W}_\gamma\overline{\Gamma}_\gamma\left|\gamma_2-\gamma_1\right|\leqslant\overline{W}_\gamma\rho_\theta^2\left|\gamma_2-\gamma_1\right|,\end{aligned}$$

where we have used the fact that $\overline{\Gamma}_\gamma = \rho_\theta^2 \left(\frac{\gamma}{\underline{\gamma}+\gamma_z}\right)^2 \leqslant \rho_\theta^2$. Repeating the same argument for $\left[\mathcal{T}^W W\right](\mu,\gamma_1) - \left[\mathcal{T}^W W\right](\mu,\gamma_2)$ to obtain the same bounds, we conclude from the definition

$$\overline{W}_\gamma = \left(1-\beta\rho_\theta^2\right)^{-1}\left(\frac{a}{2\underline{\gamma}^2}e^{-a\underline{\mu}+\frac{a^2}{2\underline{\gamma}}} + \beta \left\|s_\gamma\right\| \overline{W}_\mu\right)$$

that Lipschitz continuity of modulus $\overline{W}_\gamma$ is preserved by the mapping $\mathcal{T}^W$.

□

Lemma 12. *Under the conditions of Proposition 3, for all $W \in \mathcal{W}$, $\mathcal{T}^W W$ is differentiable.*

Proof. Differentiability in μ is immediate after proceeding to the change of variable $\mu' = \rho_\theta \mu + s\left(n^{SP},\gamma\right)\varepsilon$. Indeed, one can rewrite

$$\begin{aligned}\left[\mathcal{T}^W W\right](\mu,\gamma) \quad = \quad & \max_{f^{SP}} \int^{f^{SP}} \left[\frac{1}{a}\left(1-e^{-a\mu+\frac{a^2}{2\gamma}}\right) - \tilde{f}\right] F\left(d\tilde{f}\right) \\ & + \beta \int \left[R(W)\right]\left(\mu',\Gamma\left(n^{SP},\gamma\right)\right) \frac{1}{\sqrt{2\pi}s\left(n^{SP},\gamma\right)} e^{-\frac{\left(\mu'-\rho_\theta\mu\right)^2}{2s\left(n^{SP},\gamma\right)^2}} d\mu'.\end{aligned}$$

W being continuous and integrable, and using the Envelope Theorem, we can differentiate under the integral sign. $\mathcal{T}^W W$ is thus differentiable in μ and its derivative is

$$W_\mu(\mu,\gamma) = e^{-a\mu+\frac{a^2}{2\gamma}} + \beta \int \left[R(W)\right]\left(\mu',\Gamma\left(n^{SP},\gamma\right)\right) \frac{\rho_\theta}{\sqrt{2\pi}s^3\left(n^{SP},\gamma\right)} \left(\mu'-\rho_\theta\mu\right) e^{-\frac{\left(\mu'-\rho_\theta\mu\right)^2}{2s\left(n^{SP},\gamma\right)^2}} d\mu'.$$

Let us then turn to the differentiability along γ. The argument is inspired from Kim (1993). Let us define $\left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma,f)$ the planner's current value evaluated at some arbitrary cutoff f as

$$\begin{aligned}\left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma,f) \quad \equiv \quad & \int^{f} \left[\frac{1}{a}\left(1-e^{-a\mu+\frac{a^2}{2\gamma}}\right) - \tilde{f}\right] F\left(d\tilde{f}\right) \\ & + \beta \int \left[R(W)\right]\left(\mu',\Gamma(n,\gamma)\right) \frac{1}{\sqrt{2\pi}s(n,\gamma)} e^{-\frac{\left(\mu'-\rho_\theta\mu\right)^2}{2s(n,\gamma)^2}} d\mu',\end{aligned}$$

where $n = F(f)$. By definition, $\left[\mathcal{T}^W W\right](\mu,\gamma) = \max_f \left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma,f)$. Fix $\mu \in \left[\underline{\mu},\overline{\mu}\right]$. Pick $\gamma_1,\gamma_2 \in \left[\underline{\gamma},\gamma\right]$ arbitrarily close, denote $f_i = \operatorname{argmax}_f \left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma_i,f)$ and $n_i = F(f_i)$ for $i=1,2$. Then, for all $f \in \mathbb{R}$, we have:

$$\left[\mathcal{T}^W W\right](\mu,\gamma_2) - \left[\mathcal{T}^W W\right](\mu,\gamma_1) \geqslant \left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma_2,f) - \left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma_1,f_1).$$

In particular, we can choose $f = \tilde{f}_2$ and $\tilde{n}_2 = F\left(\tilde{f}_2\right)$ to be such that $\Gamma(n_1,\gamma_1) = \Gamma(\tilde{n}_2,\gamma_2)$, that is to say $\tilde{f}_2 = F^{-1}\left(n_1 - \frac{\gamma_2-\gamma_1}{\gamma_x}\right)$, which is well defined for γ_1 and γ_2 close enough. We then obtain the following lower bound:

$$\begin{aligned}\left[\mathcal{T}^W W\right](\mu,\gamma_2)-\left[\mathcal{T}^W W\right](\mu,\gamma_1) &\geqslant \Psi(\mu,\gamma_1,n_1,\Delta)-\Psi(\mu,\gamma_1,n_1,0)\\ &\quad +\Xi(\mu,\gamma_1,n_1,\Delta)-\Xi(\mu,\gamma_1,n_1,0),\end{aligned}$$

where $\Delta=\gamma_2-\gamma_1$ and

$$\begin{aligned}\Upsilon(\mu,\gamma,n,\Delta) &\equiv \int^{F^{-1}\left(n-\frac{\Delta}{\gamma_x}\right)}\left[\frac{1}{a}\left(1-e^{-a\mu+\frac{a^2}{2(\gamma+\Delta)}}\right)-f\right]F(df)\\ \Xi(\mu,\gamma,n,\Delta) &\equiv \beta\int\left[R(W)\right]\left(\mu',\Gamma(n,\gamma)\right)\frac{1}{\sqrt{2\pi}s\left(n-\frac{\Delta}{\gamma_x},\gamma+\Delta\right)}e^{-\frac{(\mu'-\rho_\theta\mu)^2}{2s\left(n-\frac{\Delta}{\gamma_x},\gamma+\Delta\right)^2}}d\mu'.\end{aligned}$$

Functions Υ and Ξ are continuously differentiable in Δ. Hence, by the Mean Value Theorem, there exists $\tilde{\Delta}_-\in[0,\Delta]$ such that

$$\left[\mathcal{T}^W W\right](\mu,\gamma_2)-\left[\mathcal{T}^W W\right](\mu,\gamma_1)\geqslant\left[\Upsilon_\Delta\left(\mu,\gamma_1,n_1,\tilde{\Delta}_-\right)+\Xi_\Delta\left(\mu,\gamma_1,n_1,\tilde{\Delta}_-\right)\right]\Delta.$$

Similarly, we can repeat this argument to obtain an upper bound. For all $f\in\mathbb{R}$,

$$\left[\mathcal{T}^W W\right](\mu,\gamma_2)-\left[\mathcal{T}^W W\right](\mu,\gamma_1) \leqslant \left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma,f_2)-\left[\tilde{\mathcal{T}}^W(W)\right](\mu,\gamma_1,f).$$

Let us choose $f=\tilde{f}_1$ to be such that $\Gamma(n_2,\gamma_2)=\Gamma\left(F\left(\tilde{f}_1\right),\gamma_1\right)$, that is to say $\tilde{f}_1=F^{-1}\left(n_2+\frac{\gamma_2-\gamma_1}{\gamma_x}\right)$. Define $\tilde{n}_1=F\left(\tilde{f}_1\right)$ the corresponding measure of investors. Then,

$$\begin{aligned}\left[\mathcal{T}^W W\right](\mu,\gamma_2)-\left[\mathcal{T}^W W\right](\mu,\gamma_1) &\leqslant \Upsilon(\mu,\gamma_2,n_2,0)-\Upsilon(\mu,\gamma_2,n_2,-\Delta)\\ &\quad +\Xi(\mu,\gamma_2,n_2,0)-\Xi(\mu,\gamma_2,n_2,-\Delta)\end{aligned}$$

with $\Delta=\gamma_2-\gamma_1$. There exists $\tilde{\Delta}_+\in[0,\Delta]$ such that

$$\left[\mathcal{T}^W W\right](\mu,\gamma_2)-\left[\mathcal{T}^W W\right](\mu,\gamma_1)\leqslant\left[\Upsilon_\Delta\left(\mu,\gamma_2,n_2,-\tilde{\Delta}_+\right)+\Upsilon_\Delta\left(\mu,\gamma_2,n_2,-\tilde{\Delta}_+\right)\right]\Delta.$$

The derivatives of functions Υ and Ξ being continuous in all their arguments and the optimal $n^{SP}=F\left(f^{SP}\right)$ being continuous by the Theorem of the Maximum, we may now take the following limit: for some $\gamma\in\left[\underline{\gamma},\overline{\gamma}\right]$, $f^{SP}=\operatorname{argmax}\left[\tilde{\mathcal{T}}^W(W)\right]\left(\mu,\gamma,\tilde{f}\right)$, $n^{SP}=F\left(f^{SP}\right)$, fix $\gamma_1=\gamma$ and take the limit $\Delta\to 0$ in $\gamma_2=\gamma+\Delta$. Then,

$$\Upsilon_\Delta\left(\mu,\gamma,n^{SP},0\right)+\Xi_\Delta\left(\mu,\gamma,n^{SP},0\right) \leqslant \lim_{\Delta\to 0}\frac{\left[\mathcal{T}^W W\right](\mu,\gamma+\Delta)-\left[\mathcal{T}^W W\right](\mu,\gamma)}{\Delta}$$

and

$$\lim_{\Delta\to 0}\frac{\left[\mathcal{T}^{W}W\right](\mu,\gamma+\Delta)-\left[\mathcal{T}^{W}W\right](\mu,\gamma)}{\Delta}\leqslant \Upsilon_{\Delta}\left(\mu,\gamma,n^{SP},0\right)+\Xi_{\Delta}\left(\mu,\gamma,n^{SP},0\right).$$

Function $\mathcal{T}^{W}W$ is thus differentiable in γ and its derivative is $\Upsilon_{\Delta}\left(\mu,\gamma,n^{SP},0\right)+\Xi_{\Delta}\left(\mu,\gamma,n^{SP},0\right)$, where

$$\begin{aligned}\Upsilon_{\Delta}\left(\mu,\gamma,n^{SP},0\right) &= -\frac{1}{\gamma_x}\frac{1}{F'\left(f^{SP}\right)}\left[\frac{1}{a}\left(1-e^{-a\mu+\frac{a^2}{2\gamma}}\right)-f^{SP}\right]+\frac{a}{2\gamma^2}e^{-a\mu+\frac{a^2}{2\gamma}}F\left(f^{SP}\right)\\ \Xi_{\Delta}\left(\mu,\gamma,n^{SP},0\right) &= \beta\int\left[R\left(W\right)\right]\left(\mu',\Gamma\left(n,\gamma\right)\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{\left(\mu'-\rho_\theta\mu\right)^2}{2s\left(n^{SP},\gamma\right)^2}}\\ &\quad\times\left(-\frac{1}{s\left(n^{SP},\gamma\right)^2}+\frac{\left(\mu'-\rho_\theta\mu\right)^2}{s\left(n^{SP},\gamma\right)^4}\right)\left(-\frac{1}{\gamma_x}s_n\left(n^{SP},\gamma\right)+s_\gamma\left(n^{SP},\gamma\right)\right)d\mu'.\end{aligned}$$

□

Lemma 13. *The social value of information is positive. In particular, the value function that corresponds to the planning problem satisfies*

$$\frac{d}{dn}\beta\mathbb{E}\left[W\left(\rho_\theta\mu+s\left(n,\gamma\right)\varepsilon,\Gamma\left(n,\gamma\right)\right)\right]\geqslant 0.$$

Proof. Since $E_0\left[u\left(\theta_t\right)\right]$ is bounded, the Principle of Optimality (Theorem 9.2 of Stokey et al. (1989)) applies and the recursive social planner problem stated in (38) is equivalent to its sequential formulation. For our purpose, it is useful to define the continuation value as

$$\begin{aligned}W^{cont}\left(\mu,\gamma\right) &= \max_{f_t^{eff}\left(\left\{X_s,Y_s\right\}_{s=0}^{t-1}\right)}\mathbb{E}_0\left\{\sum_{t=1}^{\infty}\beta^t F\left(f_t^{eff}\right)\mathbb{E}_t\left[u\left(\theta_t\right)\mid\left\{X_s,Y_s\right\}_{s=0}^{t-1}\right]\right\} \qquad (44)\\ \text{s.t.}\quad &\forall t\geqslant 0,\ \theta_{t+1}=\rho_\theta\theta_t+\varepsilon_t^\theta\\ &X_t=\theta_t+\varepsilon_t^X, Y_t=\theta_t+\varepsilon_t^Y\\ &\theta_0\sim\mathcal{N}\left(\mu,\gamma^{-1}\right),\end{aligned}$$

which, by the Principle of Optimality, verifies $W^{cont}\left(\mu,\gamma\right)=\beta\mathbb{E}\left[W\left(\mu',\gamma'\right)\mid\mu,\gamma\right]$. We also define the augmented problem, which corresponds to the sequential planning problem when the planner is initially endowed with an additional signal $v=\theta_0+\varepsilon_0^v$ where $\varepsilon_0^v\sim\mathcal{N}\left(0,\left(\Delta n\gamma_x\right)^{-1}\right)$ for some $\Delta n>0$, as

$$\begin{aligned}\tilde{W}^{cont,\Delta n}\left(\mu,\gamma,v\right) &= \max_{f_t^{eff}\left(\left\{X_s,Y_s\right\}_{s=0}^{t-1},v_0\right)}\mathbb{E}_0\left\{\sum_{t=1}^{\infty}\beta^t F\left(f_t^{eff}\right)\mathbb{E}_t\left[u\left(\theta_t\right)\mid\left\{X_s,Y_s\right\}_{s=0}^{t-1},v_0\right]\right\} \qquad (45)\\ \text{s.t.}\quad &\forall t,\ \theta_{t+1}=\rho_\theta\theta_t+\varepsilon_t^\theta\\ &X_t=\theta_t+\varepsilon_t^X, Y_t=\theta_t+\varepsilon_t^Y\\ &\theta_0\sim\mathcal{N}\left(\mu,\gamma^{-1}\right), v=\theta_0+\varepsilon_0^v.\end{aligned}$$

Note first that any feasible strategy f^{SP} in the planning problem (44) is feasible in the augmented problem (45) by defining

$$f_t^{eff,v}\left(\{X_s,Y_s\}_{s=0}^{t-1},v\right)\equiv f_t^{eff}\left(\{X_s,Y_s\}_{s=0}^{t-1}\right),\forall t,v,\{X_s,Y_s\}_{s=0}^{t-1}.$$

Hence, applying the law of iterated expectation, the value functions satisfy $\tilde{W}^{cont,\Delta n}(\mu,\gamma,v)\geqslant W^{cont}(\mu,\gamma)$ and the marginal social value of a signal is positive. Note, in addition, that the signal v only enters the problem (45) through the expectation operator $\mathbb{E}_t[\cdot]$. In particular, using Bayes' formula, $\forall t$,

$$\mathbb{E}_t\left[u(\theta_t)\mid\theta_0\sim\mathcal{N}\left(\mu,\gamma^{-1}\right),v,\{X_s,Y_s\}_{s=0}^{t-1}\right]=\mathbb{E}_t\left[u(\theta_t)\mid\theta_0\sim\mathcal{N}\left(\frac{\gamma\mu+\Delta n\gamma_x v}{\gamma+\Delta n\gamma_x},(\gamma+\Delta n\gamma_x)^{-1}\right),\{X_s,Y_s\}_{s=0}^{t-1}\right],$$

which implies that $\tilde{W}^{cont,\Delta n}(\mu,\gamma,v)=W^{cont}\left(\frac{\gamma\mu+\Delta n\gamma_x v}{\gamma+\Delta n\gamma_x},\gamma+\Delta n\gamma_x\right),\forall\mu,\gamma,v$. To conclude, let us return to the value the value $\frac{d}{dn}\beta\mathbb{E}[W\mid\mu,\gamma]$. Note that since W is differentiable in μ and γ from Lemma (12), this derivative exists and is well defined. We now note that, by definition, $\frac{d}{dn}\beta\mathbb{E}[W\mid\mu,\gamma]$ is the value of adding one marginal signal v about θ of precision "$\gamma_x dn$" at the end of a period. By construction, we have

$$\begin{aligned}\frac{d}{dn}\beta\mathbb{E}\left[W\left(\mu',\gamma'\right)\mid\mu,\gamma\right] &= \lim_{\Delta n\to 0}(\Delta n)^{-1}\mathbb{E}\left[W^{cont}\left(\frac{\gamma\mu+\Delta n\gamma_x v}{\gamma+\Delta n\gamma_x},\gamma+\Delta n\gamma_x\right)-W(\mu,\gamma)\mid\mu,\gamma\right]\\ &= \lim_{\Delta n\to 0}(\Delta n)^{-1}\mathbb{E}\left[\tilde{W}^{cont,\Delta n}(\mu,\gamma,v)-W(\mu,\gamma)\mid\mu,\gamma\right]\geqslant 0.\end{aligned}$$

□

Lemma 14. *Let $W\in\mathcal{W}$ be the unique solution to the planner's problem and f^{SP} its optimal cutoff. Let $V^{\tau}(\mu,\gamma,f)$ be the value function of a firm facing investment subsidy $\tau(\mu,\gamma)$ and tax $T(\mu,\gamma)$ such that*

$$V^{\tau}(\mu,\gamma,f)=\max\left\{\mathbb{E}[u(\theta)\mid\mu,\gamma]-f+\tau(\mu,\gamma),\beta\mathbb{E}\left[V^{\tau}\left(\mu',\gamma',f'\right)\right]\right\}-T(\mu,\gamma). \tag{46}$$

Then, the subsidy τ and tax T implicitly defined by

$$\begin{aligned}\tau(\mu,\gamma) &= \beta\frac{d}{dn}\mathbb{E}\left[W\left(\rho_\theta\mu+s\left(n^{SP},\gamma\right)\varepsilon,\Gamma\left(n^{SP},\gamma\right)\right)\right]\\ &\quad+\beta\iint V^{\tau}\left(\rho_\theta\mu+s\left(n^{SP},\gamma\right)\varepsilon,\Gamma\left(n^{SP},\gamma\right),f'\right)\Phi(d\varepsilon)F\left(df'\right),\end{aligned} \tag{47}$$

where $n^{SP}=F\left(f^{SP}\right)$ and $T(\mu,\gamma)=F\left(f^{SP}(\mu,\gamma)\right)\tau(\mu,\gamma)$, exist and implement the planner's allocation.

Proof. Expression (47) is a functional equation in τ because the value function V^{τ} depends implicitly on τ. To show that this functional equation has a solution, we define the mapping $\mathcal{T}^{\tau}$ corresponding to (46) on the set of continuous and bounded functions $B(\mathcal{S}\times\mathbb{R})$ to itself such that for all $(\mu,\gamma,f)\in\mathcal{S}\times\mathbb{R}$ and $V\in B(\mathcal{S}\times\mathbb{R})$,

$$\begin{aligned}
\left[\mathcal{T}^{\tau}V\right](\mu,\gamma,f) &= \max\left\{\mathbb{E}\left[u(\theta)\mid\mu,\gamma\right]-f+\tau(\mu,\gamma),\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)\right]\right\}-T(\mu,\gamma)\\
\text{s.t.} \quad & \tau(\mu,\gamma)=\beta\frac{d}{dn}\mathbb{E}\left[W\left(\mu',\gamma'\right)\right]+\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)\right]\\
& T(\mu,\gamma)=n^{SP}(\mu,\gamma)\,\tau(\mu,\gamma)\\
& \mu'=\rho_{\theta}\mu+s\left(n^{SP}(\mu,\gamma),\gamma\right)\varepsilon,\ \varepsilon\in\mathcal{N}(0,1),\\
& \gamma'=\Gamma\left(n^{SP}(\mu,\gamma),\gamma\right).
\end{aligned}$$

Note that the mapping $\mathcal{T}^{\tau}$ satisfies Blackwell's sufficient conditions.

1. *Monotonicity*: note that

$$\begin{aligned}
\left[\mathcal{T}^{\tau}V\right](\mu,\gamma,f) = \max\Bigg\{&\mathbb{E}\left[u(\theta)\mid\mu,\gamma\right]-f+\left(1-n^{SP}\right)\times\\
&\left[\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)\right]+\beta\frac{d}{dn}\mathbb{E}\left[W\left(\mu',\gamma'\right)\right]\right],\\
&\left(1-n^{SP}\right)\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)\right]-n^{SP}\beta\frac{d}{dn}\mathbb{E}\left[W\left(\mu',\gamma'\right)\right]\Bigg\}.
\end{aligned}$$

   Hence, for $V_1\leqslant V_2\in B(\mathcal{S}\times\mathbb{R})$, $\mathcal{T}^{\tau}V_2\geqslant\mathcal{T}^{\tau}V_1$.

2. *Discounting*: for $a\in\mathbb{R}$,

$$\begin{aligned}
\left[\mathcal{T}^{\tau}(V+a)\right](\mu,\gamma,f) &= \max\Bigg\{\mathbb{E}\left[u(\theta)\mid\mu,\gamma\right]-f+\left(1-n^{SP}\right)\times\\
&\quad\left[\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)+a\right]+\beta\frac{d}{dn}\mathbb{E}\left[W\left(\mu',\gamma'\right)\right]\right],\\
&\quad\left(1-n^{SP}\right)\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)+a\right]-n^{SP}\beta\frac{d}{dn}\mathbb{E}\left[W\left(\mu',\gamma'\right)\right]\Bigg\}\\
&\leqslant \left[\mathcal{T}^{\tau}V\right](\mu,\gamma,f)+\beta a.
\end{aligned}$$

The mapping $\mathcal{T}^{\tau}$ thus defines a contraction on $B(\mathcal{S}\times\mathbb{R})$. Denote V^{τ} its unique fixed point. The subsidy $\tau(\mu,\gamma)=\beta\frac{d}{dn}\mathbb{E}_{n^{SP}}\left[W\left(\mu',\gamma'\right)\right]+\beta\mathbb{E}_{n^{SP}}\left[V^{\tau}\left(\mu',\gamma',f'\right)\right]$ and the tax $T(\mu,\gamma)=n^{SP}(\mu,\gamma)\,\tau(\mu,\gamma)$ implement the efficient cutoff f^{SP}. Indeed, the maximization yields the following decision: invest if and only if

$$f \leq \mathbb{E}\left[u(\theta)\mid\mu,\gamma\right]+\tau(\mu,\gamma)-\beta\mathbb{E}\left[V\left(\mu',\gamma',f'\right)\right]=\mathbb{E}\left[u(\theta)\mid\mu,\gamma\right]+\beta\frac{d}{dn}\mathbb{E}\left[W\left(\mu',\gamma'\right)\right],$$

which coincides with the efficient cutoff f^{SP}. □

G.5 Lemmas for Proposition 4

Lemma 15. *Under the conditions of Proposition 4, for γ_x small enough, the planner's optimal cutoff f^{SP} is continuous and strictly increasing in μ and γ such that*

$$\forall \quad \underline{\mu} \leqslant \mu_1 \leqslant \mu_2 \leqslant \overline{\mu}, \quad f^{SP}(\mu_2, \gamma) - f^{SP}(\mu_1, \gamma) \geqslant \frac{1}{3} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}} (\mu_2 - \mu_1),$$
$$\forall \quad \underline{\gamma} \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \overline{\gamma}, \quad f^{SP}(\mu, \gamma_2) - f^{SP}(\mu, \gamma_1) \geqslant \frac{a}{3\overline{\gamma}^2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}} (\gamma_2 - \gamma_1).$$

Proof. Expanding the derivative in equation (39) (Proposition 3), the planner's optimal cutoff solves the implicit equation

$$\begin{aligned} f^{SP}(\mu, \gamma) &= \frac{1}{a}\left(1 - e^{-a\mu + \frac{a^2}{2\gamma}}\right) + \beta \int \varepsilon s_n\left(n^{SP}, \gamma\right) W_\mu\left(\rho_\theta \mu + s\left(n^{SP}, \gamma\right)\varepsilon, \Gamma\left(n^{SP}, \gamma\right)\right) \Phi(d\varepsilon) \\ &\quad + \beta \int \Gamma_n\left(n^{SP}, \gamma\right) W_\gamma\left(\rho_\theta \mu + s\left(n^{SP}, \gamma\right)\varepsilon, \Gamma\left(n^{SP}, \gamma\right)\right) \Phi(d\varepsilon) \end{aligned}$$

where $n^{SP} = F\left(f^{SP}\right)$. To analyze the behavior of the optimal cutoff, define the function $\vartheta(\mu, \gamma, f)$ as

$$\begin{aligned} \vartheta(\mu, \gamma, f) &\equiv \frac{1}{a}\left(1 - e^{-a\mu + \frac{a^2}{2\gamma}}\right) + \beta \int \varepsilon s_n(F(f), \gamma) W_\mu(\rho_\theta \mu + s(F(f), \gamma)\varepsilon, \Gamma(F(f), \gamma)) \Phi(d\varepsilon) \\ &\quad + \beta \int \Gamma_n(F(f), \gamma) W_\gamma(\rho_\theta \mu + s(F(f), \gamma)\varepsilon, \Gamma(F(f), \gamma)) \Phi(d\varepsilon) - f. \end{aligned}$$

By definition, $f^{SP}(\mu, \gamma)$ is a root of function ϑ, i.e., $\vartheta\left(\mu, \gamma, f^{SP}(\mu, \gamma)\right) = 0$. To lighten the notation, define the functions A and B as

$$\begin{aligned} A(\mu, \gamma, f) &\equiv \int \varepsilon \frac{\rho_\theta}{2(\gamma + \gamma_z + n\gamma_x)^2} \left(\frac{1}{\gamma} - \frac{1}{\gamma + \gamma_z + n\gamma_x}\right)^{-\frac{1}{2}} W_\mu(\rho_\theta \mu + s(F(f), \gamma)\varepsilon, \Gamma(F(f), \gamma)) \Phi(d\varepsilon) \\ B(\mu, \gamma, f) &\equiv \int \frac{\rho_\theta^2 \Gamma(n, \gamma)^2}{(\gamma + \gamma_z + n\gamma_x)^2} W_\gamma(\rho_\theta \mu + s(F(f), \gamma)\varepsilon, \Gamma(F(f), \gamma)) \Phi(d\varepsilon), \end{aligned}$$

where we have substituted for the expressions of s_n and Γ_n, so that

$$\vartheta(\mu, \gamma, f) = \frac{1}{a}\left(1 - e^{-a\mu + \frac{a^2}{2\gamma}}\right) - f + \beta\gamma_x [A(\mu, \gamma, f) + B(\mu, \gamma, f)].$$

The derivatives of W, provided in Lemma 12, are algebraic combinations of continuously differentiable functions with bounded derivatives on a compact and the Lipschitz function W. They are therefore Lipschitz in μ and γ. Functions A and B are themselves compositions and multiplications of continuously differentiable functions on a compact with W_μ and W_γ. Hence, they are Lipschitz continuous in μ, γ and f. Denote their moduli of Lipschitz continuity by $\left(\overline{A}_\mu, \overline{A}_\gamma, \overline{A}_f\right)$ and $\left(\overline{B}_\mu, \overline{B}_\gamma, \overline{B}_f\right)$. By choosing γ_x sufficiently small that $\gamma_x \beta \overline{A}_f \leqslant \frac{1}{4}$ and $\gamma_x \beta \overline{B}_f \leqslant \frac{1}{4}$, we thus have for all $f_1 \leqslant f_2$,

$$-\frac{3}{2}(f_2 - f_1) \leqslant \vartheta(\mu, \gamma, f_2) - \vartheta(\mu, \gamma, f_1) \leqslant -\frac{1}{2}(f_2 - f_1).$$

Similarly, choosing γ_x small enough that $\gamma_x \beta \overline{A}_\mu \leqslant \frac{1}{4} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}}$, $\gamma_x \beta \overline{B}_\mu \leqslant \frac{1}{4}\frac{1}{a} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}}$, $\gamma_x \beta \overline{A}_\gamma \leqslant \frac{1}{4}\frac{a}{2\overline{\gamma}^2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}}$ and $\gamma_x \beta \overline{B}_\gamma \leqslant \frac{1}{4}\frac{a}{2\overline{\gamma}^2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}}$, we have

$$\forall \quad \underline{\mu} \leqslant \mu_1 \leqslant \mu_2 \leqslant \overline{\mu}, \quad \vartheta(\mu_2, \gamma, f) - \vartheta(\mu_1, \gamma, f) \geqslant \frac{1}{2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}} (\mu_2 - \mu_1)$$
$$\forall \quad \underline{\gamma} \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \overline{\gamma}, \quad \vartheta(\mu, \gamma_2, f) - \vartheta(\mu, \gamma_1, f) \geqslant \frac{1}{2}\frac{a}{2\overline{\gamma}^2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}} (\gamma_2 - \gamma_1).$$

Hence, by the Implicit Function Theorem applied to Lipschitz functions (Dontchev and Rockafellar, 2009), the planner's optimal cutoff $f^{SP}(\mu, \gamma)$ is Lipschitz continuous and such that for γ_x small enough

$$\forall \quad \underline{\mu} \leqslant \mu_1 \leqslant \mu_2 \leqslant \overline{\mu}, \quad f^{SP}(\mu_2, \gamma) - f^{SP}(\mu_1, \gamma) \geqslant -\frac{\frac{1}{2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}}}{-\frac{3}{2}} (\mu_2 - \mu_1)$$
$$\forall \quad \underline{\gamma} \leqslant \gamma_1 \leqslant \gamma_2 \leqslant \overline{\gamma}, \quad f^{SP}(\mu, \gamma_2) - f^{SP}(\mu, \gamma_1) \geqslant -\frac{\frac{1}{2}\frac{a}{2\overline{\gamma}^2} e^{-a\overline{\mu} + \frac{a^2}{2\overline{\gamma}}}}{-\frac{3}{2}} (\gamma_2 - \gamma_1).$$

□