Essays in Monetary and Financial Macroeconomics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von

Andrey Alexandrov

im Frühjahrs-/Sommersemester 2021

Abteilungssprecher	Prof. Volker Nocke, PhD
Referent	Prof. Klaus Adam, PhD
Koreferent	Prof. John Geanakoplos, PhD

Tag der Verteidigung	2. August 2021

Acknowledgements

I would like to thank my advisor, Klaus Adam, for his invaluable guidance. I am grateful to him for the continuous support throughout all stages of my PhD, for the numerous suggestions and discussions that have greatly improved my dissertation, and for introducing me to the world of academia.

The second chapter of this dissertation was inspired by the work of John Geanakoplos, and I am grateful to him for encouraging me to venture into a new topic and for his feedback and advice on this path.

This dissertation has benefited from countless interactions with other researchers, and I am especially grateful to Matthias Meier, Michèle Tertilt and Aleh Tsyvinski for their feedback and advice. I would also like to thank Antoine Camous, Antonio Ciccone, Andreas Gulyas, Lukas Henkel, Ilse Lindenlaub, Michael Peters, Herakles Polemarchakis, Tony Smith, as well as participants of the Mannheim Macro Seminar and Klaus Adam's Colloquium for their comments and suggestions.

Thank you Daisuke Adachi, Marianne Bernatzky Koehli, Enrico Camarda, Luis Carvalho Monteiro, Moritz Drechsel-Grau, Chuan Du, Tsung-Hsien Li, Alexander Matusche, Tomasz Olma, Timo Reinelt, Mykola Ryzhenkov, Suk Joon Son and Zhengren Zhu for the good time and fun I had during the PhD. I am especially grateful to Fabian Greimel and Claudio Kretz for the great discussions that we had in our dissertation group.

I want to thank Kerry Bowman, Marion Lehnert, Caroline Mohr, Susanne Putz and Sylvia Rosenkranz for their administrative support and for making my life a lot easier. I am grateful to the University of Mannheim, the German Academic Exchange Service (DAAD), the foundation Geld & Währung, and the German Research Foundation (DFG) for financial support, and to the state of Baden-Württemberg for computational resources through bwHPC.

And finally, I would like to thank my family. I am grateful to my parents, who not only supported me my entire life, but also encouraged and celebrated all my academic endeavors. I am grateful to my wife for her support and patience, for proof-reading countless versions of this dissertation, and for turning the PhD into an exciting and enjoyable journey.

Contents

Preface

This dissertation consists of two chapters. In the following I provide a brief introduction to each of them.

In the first chapter I study the effects of trend inflation on aggregate price and output dynamics and their implications for the effectiveness of monetary stabilization policy. This chapter relates to the ongoing discussion on whether central banks should raise their inflation targets and achieve higher levels of trend inflation. In 2020 both the Federal Reserve and the European Central Bank have signaled changes in monetary policy that are expected to result in higher levels of trend inflation. This, in turn, would certainly change the price setting behaviour of firms, which is central for the transmission of monetary policy. To understand these effects, I address the following question – how would a higher level of trend inflation affect the price setting behaviour of firms and what are the consequences for aggregate price and output dynamics?

To answer this question, I use the standard heterogeneous agents model of price setting in continuous time and characterize these effects analytically. The key result of this chapter is that trend inflation affects responses to expansionary and contractionary shocks asymmetrically. In particular, it amplifies price responses and mitigates output responses to expansionary shocks, whereas the effects are reversed for contractionary shocks. Furthermore, I show that under positive trend inflation, sufficiently large expansionary monetary shocks lead to a decline in output. Counterintuitively, the monetary authority tries to stimulate the economy, but achieves quite the opposite – a contraction in output.

I find that these model predictions are empirically supported by U.S. sectoral data and calibrate a quantitative version of the model to show why the asymmetry matters for monetary stabilization policy. I consider an adverse markup shock that increases desired markups of firms and leads to higher prices and lower output. This creates a stabilization trade-off for the monetary authority because it can not stabilize prices and output simultaneously. I show that raising trend inflation from 2% to 4% has two adverse effects. First, it amplifies the economy's response to the markup shock and, second, it worsens the trade-off between output and price stability. Higher trend inflation reduces the ability of the monetary authority to stabilize the economy because prices become stickier exactly when flexibility is needed, whereas they become more flexible exactly when rigidity is needed.

In the second chapter I study the determinants of leverage and asset prices in a general equilibrium model with incomplete markets and collateral requirements. Secured (collateralized) loans are a common type of borrowing, in which borrower pledges an asset as collateral to ensure lender against potential default. In many cases, agents borrow to purchase an asset and use that asset to secure the loan. A key statistic that affects the ability of borrowers to purchase assets is leverage, which is the ratio between the value of an asset and the agent's down payment. Leverage on secured loans experienced violent fluctuations around the Great Recession. The aim of this chapter is to provide new theoretical insights into forces and mechanisms responsible for such drastic movements.

To understand the determinants of leverage, I setup a general equilibrium model with an endogenous leverage constraint. The agents disagree about the distribution of the risky asset payoff, meaning that some agents are more optimistic than the others. In equilibrium, optimists would like to buy the risky asset and to borrow from the more pessimistic agents to increase their asset purchases. However, borrowing needs to be collateralized by the risky asset, the value of which is different for borrowers and lenders. This discrepancy creates an endogenous leverage constraint for the borrowers. The main contribution of this chapter is that it features a continuum of agent types and a continuum of possible future asset payoffs, in contrast to the previous literature. Studying two types of continua provides new insights, unattainable in simpler environments.

I show that agents sort into three categories: pessimists who buy the safe asset, optimists who buy the risky asset leveraged, and agents with moderate optimism who lend to the optimists. Each borrower-lender pair has a unique contract in terms of interest rate and leverage. A new analytic result shows that in equilibrium only risky borrowing contracts are traded, meaning that if the asset payoff is low enough, all contracts default. In addition, leverage and the asset price are decoupled in equilibrium and can move independently. The asset price depends on the total mass of market participants, whereas leverage – on how these agents split into borrowers and lenders. Numerical simulations suggest that leverage, unlike the price of the asset, is not sensitive to changes in the average optimism but drops significantly if uncertainty rises. The reason is that higher uncertainty scares lenders away. Some of them decide not to participate in the market at all and switch from risky lending to buying a safe asset. Others decide to switch sides of the market and buy the asset on margin instead.

While the two chapters address questions in different areas of economics, they share a common theme: I study how the interaction between heterogeneity and micro-level frictions shapes aggregate outcomes. In the first chapter, heterogeneity is driven by idiosyncratic shocks, which give rise to a cross-sectional distribution of firms, and the friction is represented by a fixed cost of price adjustment. Because of the adjustment cost, trend inflation incentivizes firms to respond differently to positive and negative shocks, which at the aggregate level is further amplified by the effect of trend inflation on the shape of the cross-sectional distribution. In the second chapter, agents have heterogeneous beliefs about the distribution of the future asset payoff, and the friction is the collateral requirement for the borrowers. Changes in fundamentals, such as aggregate optimism or uncertainty, affect the endogenous leverage constraint and alter the identities of borrowers, lenders and safe asset investors, which ultimately determine the aggregate leverage and the asset price.

Chapter 1

The Effects of Trend Inflation on Aggregate Dynamics and Monetary Stabilization

1.1 Introduction

Over the last decade there has been a discussion on whether central banks should raise their inflation targets and achieve higher levels of trend inflation. In August 2020 the Federal Reserve announced a major change to its policy, which would allow inflation to stay above the 2% target after a period of below-target inflation, whereas previously the Federal Reserve would not have tolerated such deviations. Just a month later, in September 2020, the European Central Bank stated that it would consider shifting to a symmetric inflation target, instead of targeting inflation rates of "below, but close to, 2%". Given the currently very low levels of inflation, both of these measures are aimed at increasing inflation expectations, which would in turn decrease expected real interest rates and stimulate consumption.

While these measures are designed to have a positive effect on the demand of households, they would also affect inflation expectations of firms. This would change the price-setting behavior of firms, which is central for the transmission of monetary policy because it determines aggregate price and output responses to shocks. In this paper I study how changes in the level of trend inflation affect the price-setting behavior of firms and the aggregate price and output dynamics. The results have important implications for the effectiveness of monetary stabilization policy and, more generally, provide new insights into aggregate dynamics in economies with lumpy adjustments.

First, I analytically characterize aggregate price and output dynamics in economies with small levels of trend inflation. I show that changes in the level of trend inflation affect aggregate responses to positive and negative shocks *asymmetrically.* A higher level of trend inflation reduces the overall potency of monetary policy to stimulate output and reverts its effects for sufficiently large shocks. I obtain novel analytic results by studying shocks of arbitrary size, in contrast to the previous literature which focused on marginal shocks. Second, I provide supporting empirical evidence for the new analytic predictions, using U.S. sectoral data. Third, I find that the effect of trend inflation is sizable in a general equilibrium model calibrated to the U.S. economy. I show that a higher inflation target impedes the ability of a monetary authority to counteract adverse shocks that move output and prices in opposite directions. Finally, many of the analytic results are of a more general interest, as they can be applied in other environments with lumpy adjustments, including models of capital and labor adjustment, and durable good consumption.

New Analytic Results. I use the workhorse menu cost model of price dynamics, in which firms face an exogenously given desired price, determined by common trend inflation (drift) and idiosyncratic shocks. Flow profit is maximized when the actual price is equal to the desired one, and price adjustment comes at a fixed cost. Because of the cost, firms keep their prices constant most of the time and adjust infrequently, which results in gaps between actual and desired prices. These price gaps are the key variable in the model and the evolution of the price gap distribution determines aggregate price and output dynamics.

Following the literature, I consider an unexpected permanent one-time monetary shock. To avoid ambiguity, I label shocks as ‘positive’ vs. ‘negative’ when referring to the intended effects of monetary policy (e.g., interest rate cuts vs. hikes). I show that the key property of trend inflation is that it affects responses to positive and negative shocks *asymmetrically.* In particular, increasing the level of trend inflation amplifies price responses to positive monetary shocks and mitigates price responses to negative monetary shocks. Since the strength of output responses is inversely related to the magnitude of aggregate price responses, the effect of trend inflation on output is reversed. With a higher level of trend inflation, output becomes less sensitive to positive monetary shocks and more sensitive to negative monetary shocks.

There are two channels through which trend inflation creates asymmetry in aggregate dynamics: the optimal policy of firms and the shape of the price gap distribution. First, if trend inflation is positive, firms expect higher prices in the future and are eager to increase them once a positive shock arrives, despite the adjustment cost. For the same reason, firms are reluctant to decrease their prices after a negative shock because it induces additional adjustment costs in the future. Second, trend inflation erodes relative prices and leads to a higher concentration of price gaps at the bottom of the price gap distribution. Thus, under positive trend inflation, positive shocks trigger more firms to adjust compared to negative shocks. Both of these channels work in the same direction and result in asymmetric aggregate price and output responses, with the degree of asymmetry depending on the level of trend inflation.

Positive trend inflation has two additional implications: price overshooting and output contractions after sufficiently large positive monetary shocks. A shock is considered ‘large’ if it forces all firms to adjust their prices. In an economy with zero trend inflation, large shocks are neutral – aggregate price responds one-to-one, and output does not move. I show that in economies with positive trend inflation, large positive shocks cause aggregate price overreaction and actually reduce output. Firms prefer to overshoot when adjusting upward, as they anticipate relative price erosion due to positive trend inflation. Overshooting at the aggregate level occurs if the mass of adjusters is sufficiently large, which highlights a special effect of the drift on aggregate responses to large shocks. In fact, a shock does not have to force all firms to adjust to have such an effect – even smaller positive shocks can cause a decline in output. Therefore, the overall ability of a monetary authority to stimulate output deteriorates as trend inflation rises: moderate shocks cause weaker responses, and larger shocks become counterproductive.

Empirical Evidence. I show that the new analytic results are supported by the data. To ensure enough variation in the level of trend inflation, I use U.S. sectoral data on the Producer Price Index (PPI) and industrial production (IP). I compute trend inflation for each sector as the average PPI growth rate and split sectors into two groups: those with trend inflation above and below the median. I then estimate non-linear impulse responses to identified monetary shocks within each group and compare the results between the two groups.

First, I find that trend inflation is strongly related to the degree of asymmetry in PPI and production responses to monetary shocks. Price responses in sectors with high trend inflation exhibit primarily positive asymmetry, i.e., prices rise more after positive shocks than they fall after negative shocks. On the contrary, the asymmetry of price responses is negative in sectors with low trend inflation. Responses of industrial production are generally negatively asymmetric, meaning that production contracts more after negative shocks than it rises after positive shocks. However, the asymmetry is much more negative in sectors with higher trend inflation, where positive shocks have almost no impact on production, and negative shocks cause substantial responses. The model does not always match the *level* of asymmetry in the data, but it correctly predicts the *relationship* between trend inflation and asymmetry. Even though the results can not be interpreted in a causal sense, they show that the model predictions are in line with the data.

Second, I find that production responses to positive shocks have an inverse U-shape, meaning that the maximum stimulative effect on production is achieved for moderate shock sizes. In addition, sufficiently large positive shocks have a reverse effect, leading to a contraction in production. As predicted by the model, these reverse effects are strongly related to the level of trend inflation: the size of a positive shock that leads to a zero production response is substantially smaller in sectors with higher trend inflation. The results suggest that monetary policy is not only less effective in stimulating output in sectors with higher trend inflation, but also has much less 'room' for doing so.

Finally, I provide evidence for the mechanism linking trend inflation and asymmetries in aggregate responses to monetary shocks. I use daily item-level price data provided by the Billion Prices Project to study the relationship between trend inflation and asymmetry in individual price adjustments. I find that a one percentage point increase in monthly trend inflation at the item level is associated with a 5% increase in the ratio between the magnitudes of positive and negative adjustments. This relationship between trend inflation and micro-level asymmetries matches the model predictions and is an important channel leading to aggregate asymmetries in responses to monetary shocks, as observed in the sectoral data.

Implications for Monetary Policy. Lastly, I show that the effects of trend inflation are sizable and relevant for policy. I embed the analytic framework into a general equilibrium model calibrated to the U.S. economy. I consider an adverse markup shock and study the ability of a policymaker to stabilize the economy in the baseline model with a 2% inflation target (trend inflation). I then compare the results with a counterfactual economy, in which the inflation target is set to 4%.

The analysis is positive, as I do not consider optimal policy, but assume a simple stabilization objective instead. A markup shock suits this purpose well, as it amplifies the inefficiency stemming from price dispersion but does not affect the efficient allocation, which rationalizes the stabilization objective. In addition, the shock increases prices and depresses consumption, introducing a trade-off for the monetary authority, as it can not stabilize consumption and prices simultaneously.

I find that raising the inflation target from 2% to 4% has two adverse effects. First, it amplifies the initial reaction to the markup shock, causing larger consumption and price deviations. Second, it worsens the stabilization trade-off. A policymaker must sacrifice more consumption when stabilizing prices and tolerate larger price deviations when stimulating consumption. These effects are due to weaker upward price rigidity and stronger downward price rigidity, caused by a higher inflation target. Increasing the inflation target leads to more price flexibility exactly when it is desirable to have rigid prices and makes prices stickier exactly when flexibility is needed. In addition, the effects of trend inflation are more pronounced for larger shocks, in accordance with the analytic results.

Relation to the Literature. The effect of drift on individual and aggregate behavior has previously drawn the attention of many researchers. Several early theoretical contributions (Sheshinski and Weiss (1977), Mankiw and Ball (1994), Tsiddon (1993)) have shown that trend inflation can affect the magnitude of individual price adjustments and the mass of adjusting firms after aggregate nominal shocks. My work closely relates to the subsequent research, which has focused on analytic characterization of aggregate dynamics in economies with lumpy adjustments (Caballero and Engel (2007), Alvarez and Lippi (2019) and Baley and Blanco (2020)). This strand of literature has either restricted its attention to marginal aggregate shocks or considered driftless economies. I contribute to the literature by simultaneously allowing for non-zero trend inflation and providing analytic results for shocks of arbitrary size. I show that the two key statistics of aggregate price and output dynamics – the impact effect and the cumulative impulse response (CIR), are both affected by trend inflation to first order, in contrast to the results obtained for marginal shocks. The analysis requires an analytic characterization of the CIR for non-zero levels of trend inflation. Alvarez et al. (2016) show that in order to compute the CIR in an economy with no drift, it is sufficient to track agents until the first time of adjustment, as subsequent paths net out to zero in expectation. I show that non-zero drift introduces an additional term, which is related to paths after the first adjustment. I propose two ways of computing the new term analytically, and the results apply to settings well beyond the scope of this paper, including other types of aggregate shocks and functions of interest.

The empirical results of this paper provide new insights into price and output responses to monetary policy shocks. Several studies have tested whether aggregate impulse responses exhibit state dependence (Lo and Piger (2005), Auerbach and Gorodnichenko (2012), Ramey and Zubairy (2014)), asymmetries with respect to positive and negative shocks (Long and Summers (1988), Cover (1992), Angrist et al. (2018)) and non-linearities with respect to the shock size (Tenreyro and Thwaites (2016), Ascari and Haber (2020)). I add to the literature by showing that both asymmetry and non-linearity of aggregate impulse responses are tightly linked to trend inflation. In addition, I provide evidence for the mechanism behind this link and show that trend inflation affects the asymmetry of price adjustments at the micro level, even if trend inflation is low. This result complements the work of Alvarez et al. (2019) who find evidence for this relationship only in a high inflation environment.

The policy implications of my results contribute to the ongoing discussion on the role of trend inflation for the effectiveness of monetary stabilization policy. The proposal of raising inflation targets to gain sufficient policy 'room' from the zero lower bound (see Blanchard et al. (2010) and Ball (2013)) drew the attention of researchers to the consequences of higher trend inflation. Ascari and Sbordone (2014) discuss the implications of a higher inflation target for price dispersion, stability of inflation expectations and macroeconomic volatility. L'Huillier and Schoenle (2020) argue that a higher inflation target increases the frequency of price adjustments, lowers the potency of monetary policy, and thus provides an effective extra room that is smaller than the nominal one. Blanco (2020) points out that a higher inflation target increases downward price rigidity, which mitigates recessions at the zero lower bound. I show that higher trend inflation has an additional adverse impact on the effectiveness of monetary policy away from the zero lower bound, particularly for shocks that introduce a trade-off between price and output stability. Although I do not study the optimal level of trend inflation, my results have implications for the normative analysis (see Coibion et al. (2012), Adam and Weber (2019), Blanco (2020), and Diercks (2017) for a review).

Adjustment costs appear in numerous economic settings, such as models of lumpy capital and labor adjustment, durable good consumption and others. In many of these models, previous studies have noted the role of drift in shaping responses to aggregate shocks. In investment models with capital adjustment costs (e.g., Khan and Thomas (2008) and Bachmann et al. (2013)), the distribution of mandated investment is highly skewed, and responses to aggregate shocks are asymmetric. In this setting, capital depreciation plays the role of drift because it erodes capital stock. Similarly, depreciation of durable goods generates asymmetry in consumption responses to fiscal stimuli over the cycle in Berger and Vavra (2015). Jaimovich and Siu (2020) show empirically that employment in routine occupations in the U.S. falls over time and predominantly during recessions, whereas employment in non-routine occupations is increasing and does not contract in recessions. These findings provide another example of the relationship between drift and cyclical behavior and are in line with my analytic results.

Structure of the paper. The next section presents the main analytical results of this paper. Section 1.3 provides empirical evidence for these new findings. Section 1.4 discusses the implications of the results for monetary stabilization policy in a calibrated general equilibrium model. Section 1.5 concludes.

1.2 Theoretical Results

I consider the simplest version of a two-sided sS model with a quadratic objective and fixed costs of adjustment. This framework serves as an approximation to numerous applications, including models of capital, labor or price adjustment, portfolio or inventory management, as well as durable good consumption. In the following, I outline the model setup and briefly review the benchmark case of a driftless economy. I then move to economies with non-zero drift and highlight the main qualitative differences.

1.2.1 Problem of a Firm

I consider a model of a firm that sets its price subject to a menu cost, given the optimal price target.[1] The instantaneous profits of the firm are given by $\pi(z) = -z^2$, where $z = \ln p - \ln p^*$ is the log deviation of the current price p from its frictionless optimum p^*. The optimal price p^* maximizes the instantaneous profits of the firm and follows a geometric Brownian motion with drift μ:

$$d \ln p^*(t) = \mu dt - \sigma dW(t)$$

where $\sigma > 0$ and $W(t)$ is a Wiener process. In this setup, the drift μ corresponds to trend inflation and is the key parameter of interest. Price adjustment comes at a fixed cost $\kappa > 0$, so that the firm keeps its price p constant most of the time and adjusts it infrequently. In the absence of price adjustment, the price gap z evolves as follows:

$$dz(t) = -\mu dt + \sigma dW(t)$$

Whenever the firm intervenes and changes its price by $\Delta \ln p$, the price gap z jumps by the same amount and in the same direction. The profits are

[1] Here, I take the optimal price as given. In a standard setting, the price target is typically determined by a markup over marginal costs, which in turn depend both on aggregate and individual states.

discounted at rate $\rho > 0$, and the firm's objective is to maximize discounted stream of profits subject to the adjustment costs it pays upon each intervention. Its problem can be formulated entirely in terms of the price gaps z, with $\{\tau_i\}_{i=1}^{\infty}$ denoting the sequence of times when the firm adjusts and $\{\Delta z_i\}_{i=1}^{\infty}$ being the sequence of adjustments:

$$v(z) = \max_{\{\tau_i, \Delta z_i\}_{i=1}^{\infty}} \mathbb{E}\left[\int_0^{\infty} e^{-\rho t}\pi(z(t))dt - \sum_{i=1}^{\infty} e^{-\rho\tau_i}\kappa \;\middle|\; z(0) = z\right]$$

$$\text{s.t.} \quad z(t) = z(0) - \mu t + \sigma W(t) + \sum_{i=1}^{N(t)} \Delta z_i$$

where $N(t)$ is the number of adjustments that occurred until t. This constitutes a standard impulse control problem, the solution to which are boundaries of inaction region $(\underline{z}, \overline{z})$ and a return point $\hat{z}$. Whenever $z(t) \in (\underline{z}, \overline{z})$, the firm keeps its current price and lets the price gap evolve stochastically. As soon as $z(t)$ reaches one of the boundaries, the firm pays an adjustment cost κ and sets $z(t) = \hat{z}$. At any intervention time τ_i, the adjustment is given by $\Delta z_i = \hat{z} - \lim_{t \uparrow \tau_i} z(t)$, where $\lim_{t \uparrow \tau_i} z(t)$ is the value of z right before the adjustment and is equal to either $\underline{z}$ or $\overline{z}$.

The value function $v(z)$ satisfies the following Hamilton–Jacobi–Bellman equation for any $z \in (\underline{z}, \overline{z})$:

$$\rho v(z) = -z^2 - \mu v'(z) + \frac{\sigma^2}{2} v''(z)$$

together with smooth pasting conditions $v'(\underline{z}) = v'(\overline{z}) = 0$, optimality of return point $v'(\hat{z}) = 0$ and boundary conditions $v(\underline{z}) = v(\overline{z}) = v(\hat{z}) - \kappa$. These conditions constitute a system of equations, which implicitly defines the solution triplet $\{\underline{z}(\mu), \hat{z}(\mu), \overline{z}(\mu)\}$. I highlight the dependence of optimal policy on trend inflation by explicitly stating μ as its argument, even though it also depends on other model parameters.

1.2.2 Aggregate Dynamics

Assume that the economy is populated by a continuum of ex-ante identical firms that face the same drift in optimal price μ but experience idiosyncratic shocks. Firms follow the same policy $\{\underline{z}(\mu), \hat{z}(\mu), \overline{z}(\mu)\}$ and the economy has a stationary distribution of price gaps z with a cumulative distribution function $F(z, \mu)$. The corresponding density is denoted by $f(z, \mu)$ and solves the following Kolmogorov forward equation:

$$0 = \mu f_z(z, \mu) + \frac{\sigma^2}{2} f_{zz}(z, \mu)$$

together with boundary conditions $f(\underline{z}(\mu), \mu) = f(\overline{z}(\mu), \mu) = 0$, unit mass condition $\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} f(z, \mu)dz = 1$ and continuity at $z = \hat{z}(\mu)$.[2]

Following the literature, I consider an unexpected permanent nominal shock that changes the optimal (log) price $\ln p^*$ by δ simultaneously for all firms. The shock shifts the stationary distribution of price gaps in the opposite direction (because $z = \ln p - \ln p^*$), as illustrated in Figure 1.2.1 for $\delta > 0$. The stationary distribution of the price gaps is depicted by the dashed blue line, whereas the solid red line shows the density immediately after the shock has arrived, but before firms responded to it. This distribution is referred to as the initial distribution and is denoted by $F_0(z, \mu)$. For this type of shock, the initial distribution is a shifted version of the stationary distribution $F(z, \mu)$, such that $F_0(z, \mu) = F(z + \delta, \mu)$.

Figure 1.2.1: Aggregate shock δ

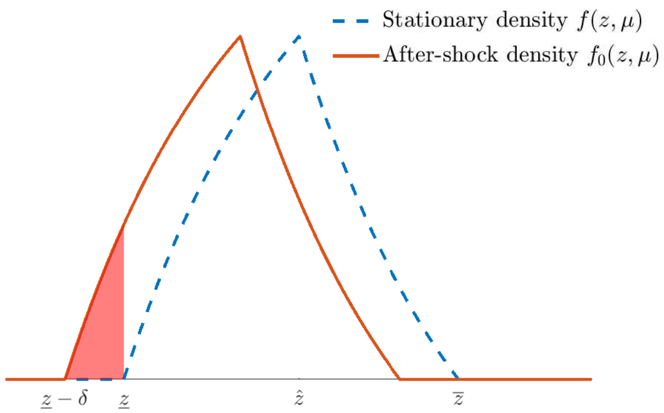


Illustration of a positive shock $\delta > 0$ in an economy with positive drift ($\mu > 0$). The dashed blue line is the stationary density of price gaps $f(z, \mu)$, whereas the solid red line is the density immediately after the shock and before firms adjust, $f_0(z, \mu)$. The shaded triangle corresponds to the mass of firms that adjust on impact.

The shaded area corresponds to the mass of firms that are pushed outside the inaction region and adjust immediately on impact. Their adjustment results in an immediate change of the aggregate price level, denoted by $\Theta(\delta, \mu)$ and commonly referred to as the impact effect. Formally, for a positive shock $\delta > 0$, $\Theta(\delta, \mu)$ is given by the following expression:

$$\Theta(\delta, \mu) = \int_{\underline{z}(\mu)-\delta}^{\underline{z}(\mu)} \big(\hat{z}(\mu) - z\big) f(z+\delta, \mu)dz \tag{1.1}$$

[2] The density $f(z, \mu)$ is non-differentiable at the return point $\hat{z}$ and the Kolmogorov forward equation does not hold at this point.

which is simply the sum of all adjustments $\big(\hat{z}(\mu)-z\big)$ weighed with the initial density. The resulting distribution of price gaps then gradually converges to the stationary one, inducing a path for aggregate variables. These dynamics are summarized in Figure 1.2.2.

Figure 1.2.2: Dynamics after an aggregate shock δ

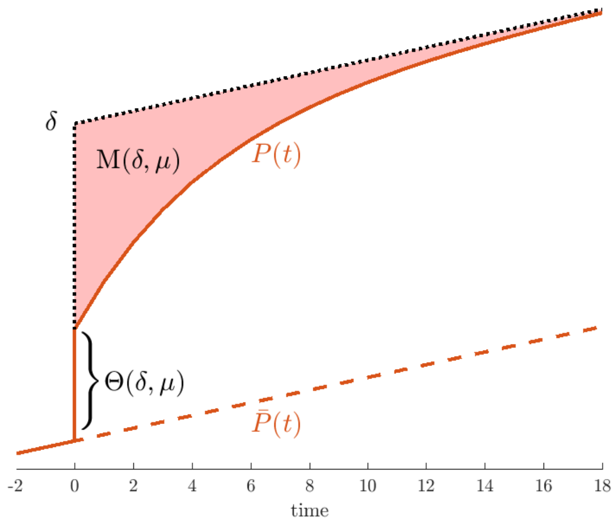


Illustration of aggregate dynamics after positive shock $\delta > 0$ in economy with positive drift ($\mu > 0$). The solid red line is the realized path of the aggregate log-price $P(t)$, the dashed red line is its hypothetical path absent of shock $\bar{P}(t)$. The initial vertical segment of $P(t)$ shows the impact effect $\Theta(\delta,\mu)$. The shaded area corresponds to the cumulative impulse response $M(\delta,\mu)$.

The solid red line shows the realized path of the aggregate log-price $P(t)$, whereas the dashed red line corresponds to its hypothetical path absent of any shock $\bar{P}(t)$. The price impulse response at any time t is given by the difference between $P(t)$ and $\bar{P}(t)$. The shock arrives at $t=0$ and triggers the impact effect $\Theta(\delta,\mu)$, given by the initial vertical jump of $P(t)$. Another statistic, commonly studied in the literature, is the cumulative impulse response (CIR), which is shown as the shaded area on the graph and denoted by $M(\delta,\mu)$:

$$M(\delta,\mu) = \int_0^\infty \big[\delta - \big(P(t) - \bar{P}(t)\big)\big]\,dt$$

This statistic summarizes the strength and speed of the price response, although in a reversed way. The stronger and faster firms react to the shock, the smaller $M(\delta,\mu)$ is. For example, if the immediate price response $\Theta(\delta,\mu)$ is equal to δ, then the shaded area in Figure 1.2.2 collapses and $M(\delta,\mu)=0$, provided that there are no further fluctuations in $P(t)$ around the trend. In addition, $M(\delta, \mu)$ is of special use in a certain class of general equilibrium models (e.g. Golosov and Lucas (2007)), as it measures the cumulative output response to a nominal shock δ. Under logarithmic preferences, the output response at any time t is given by $Y(t) = \delta - \big(P(t) - \bar{P}(t)\big)$, meaning that output absorbs the part of the shock that was not captured by the price response.[3] Cumulating output responses over time recovers the expression for $M(\delta, \mu)$, which provides an immediate mapping from price to output responses and characterizes the real effects of nominal shocks.

Because aggregate price dynamics are determined by the dynamics of the price gap distribution, one can express $M(\delta, \mu)$ in the following way:[4]

$$M(\delta, \mu) = -\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \mathbb{E}\left(\int_0^\infty \big(z(t) - \bar{x}(\mu)\big)dt \;\middle|\; z(0) = z\right) dF_\delta(z, \mu)$$

where $\bar{x}(\mu)$ is the average gap in the steady state $\Big(\bar{x}(\mu) = \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} z dF(z, \mu)\Big)$. The outer integrand is the expectation of the cumulated deviations of the price gap $z(t)$ from its steady state average $\bar{x}(\mu)$, given a particular starting value $z(0) = z$. The integrand is then averaged across all starting values z, using the distribution of price gaps immediately after the shock has arrived and firms outside of the inaction region have adjusted, which is denoted by $F_\delta(z, \mu)$. This distribution is equal to the stationary one, shifted by δ and truncated to the inaction region, together with a mass point at $\hat{z}(\mu)$ due to a positive mass of firms that adjust on impact.

There are several limitations of this framework. First, the quadratic profit function serves as a second-order approximation to a more general one, e.g., the one resulting from a CES demand function. Second, I ignore any general equilibrium feedback effects from aggregate dynamics to the optimal policy of firms to ensure that firms follow the steady state policy along the transition path.[5] Both assumptions are crucial for analytic tractability and are relaxed in Section 4, where I calibrate a general equilibrium model to the U.S. economy.

The primary interest of this paper is the sensitivity of the impact price effect $\Theta(\delta, \mu)$ and the cumulative output response $M(\delta, \mu)$ to changes in trend inflation μ, particularly for shocks that are not marginal. I briefly review the

[3] Relaxing logarithmic preferences to a more general case of CES preferences makes output responses proportional to $\delta - \big(P(t) - \bar{P}(t)\big)$.

[4] See Appendix A.1.1 for details.

[5] Alvarez and Lippi (2014) show that such a setting, these effects are of second order only.

main properties of $\Theta(\delta, \mu)$ and $M(\delta, \mu)$ in a benchmark driftless setting, and then discuss economies with non-zero drift.

1.2.3 Driftless Benchmark

Driftless economies have been well studied in the literature and serve as an important benchmark for economies with small drift. In a recent study, Alvarez and Lippi (2019) characterize the entire impulse response to any initial disturbance for economies without drift, whereas full characterization with non-zero drift is still a challenge. To allow for comparability between the setups, I keep the focus on the impact effect and the cumulative impulse response, and review their main properties in economies without drift.

The absence of drift in the optimal price coupled with a quadratic profit function results in a symmetric optimal policy $\{\underline{z}(0), \hat{z}(0), \overline{z}(0)\} = \{-\overline{z}_0, 0, \overline{z}_0\}$. The return point $\hat{z}(0)$ is set to zero and the lower boundary of inaction region $\underline{z}(0)$ is the negative of the upper boundary $\overline{z}(0)$, denoted by $\overline{z}_0$ to ease notation. Stationary density $f(z, 0)$ becomes a piecewise linear function with a kink at zero. Solving for the impact effect $\Theta(\delta, 0)$ of a positive shock $\delta > 0$ yields the following result (derivation is provided in Appendix A.1.2):

$$\Theta(\delta, 0) = \begin{cases} \frac{1}{6\overline{z}_0^2}\delta^2(3\overline{z}_0 + \delta), & \text{for } \delta < \overline{z}_0 \\ \frac{1}{6\overline{z}_0^2}\left[\delta(6\overline{z}_0^2 + 3\delta\overline{z}_0 - \delta^2) - 4\overline{z}_0^3\right], & \text{for } \delta \in [\overline{z}_0, 2\overline{z}_0) \\ \delta, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}$$

While Alvarez and Lippi (2014) characterize $\Theta(\delta, 0)$ given small ($\delta \leq \overline{z}_0$) and large ($\delta \geq 2\overline{z}_0$) values of the shock, I also derive an expression for intermediate values. Three key features of impact effect under zero drift should be highlighted. First, due to symmetries in optimal policy and stationary density, the impact effect is symmetric for positive and negative shocks, i.e., $\Theta(-\delta, 0) = -\Theta(\delta, 0)$. Second, the impact response of aggregate price never exceeds the shock: $|\Theta(\delta, 0)| \leq |\delta|$ for all δ. Third, when a shock is large ($\delta \geq 2\overline{z}_0$), the price level responds one-to-one, meaning that $\Theta(\delta, 0) = \delta$. In general, a shock is considered 'large' if it pushes all firms outside of the inaction region, forcing all of them to adjust. Therefore, the shock must be larger than the width of the inaction region $\overline{z}(\mu) - \underline{z}(\mu)$, which in the driftless case is equal to $2\overline{z}_0$. Both the average size of price adjustments $\mathbb{E}(|\Delta \ln p|)$ and the standard deviation of adjustments $Std(\Delta \ln p)$ are equal to $\overline{z}_0$, so that δ is large if it is twice as big as the average adjustment size or exceeds two standard deviations.

Alvarez et al. (2016) show that to compute the cumulative output response $M(\delta, 0)$, one does not have to consider the entire path of price gap deviations, as it is enough to keep track of each firm until the first adjustment. Because of zero drift, the expected price gap deviation is always zero after the first adjustment. Furthermore, the steady state average gap $\bar{x}(0)$ is also zero, which gives the following expression for the CIR:

$$M(\delta, 0) = -\int_{-\bar{z}_0}^{\bar{z}_0} \mathbb{E}\left(\int_0^{\tau} z(t)dt \mid z(0) = z \right) dF_\delta(z, 0)$$

where τ is the first time of adjustment. I provide an expression for $M(\delta, 0)$ in Appendix A.1.2 and briefly review its main properties below.

Identical to the impact effect, there are three key features to note: (1) CIR is symmetric around zero in the sense that $M(-\delta, 0) = -M(\delta, 0)$; (2) The cumulative output response is non-negative for all $\delta > 0$, meaning that positive nominal shocks either increase output or are neutral; (3) Large shocks are neutral ($M(\delta, 0) = 0$), because aggregate price adjusts one-to-one to these shocks on impact and $\Theta(\delta, 0) = \delta$.

As I show in subsequent sections, none of the main properties of the impact effect and the cumulative output response are valid in economies with non-zero drift.

1.2.4 Introducing Drift

I now study economies with non-zero drift: $\mu \neq 0$. The optimal inaction region of a firm is no longer symmetric and the return point is not zero. Given that the problem is well characterized for zero drift, I consider a first-order approximation of the key statistics around the zero drift point:[6]

$$\Theta(\delta, \mu) = \Theta(\delta, 0) + \frac{\partial \Theta(\delta, 0)}{\partial \mu}\mu + o(\mu^2)$$
$$M(\delta, \mu) = M(\delta, 0) + \frac{\partial M(\delta, 0)}{\partial \mu}\mu + o(\mu^2)$$

This approach is novel, as I compute the first derivatives of aggregate responses with respect to the drift μ for shocks of any size. To date, the literature has only considered the effect of drift on responses to marginal shocks, given by cross-derivatives $\frac{\partial^2 \Theta(\delta,\mu)}{\partial\delta\partial\mu}\big|_{\delta=0,\mu=0}$ and $\frac{\partial^2 M(\delta,\mu)}{\partial\delta\partial\mu}\big|_{\delta=0,\mu=0}$. Alvarez et al. (2016) show that these cross-derivatives are equal to zero due to the symmetry properties of the model and the assumed differentiability of

[6] I use short-hand notation $\frac{\partial X(\delta,0)}{\partial \mu}$ for $\frac{\partial X(\delta,\mu)}{\partial \mu}\big|_{\mu=0}$.

$\Theta(\delta,\mu)$ and $M(\delta,\mu)$ with respect to μ. This result does not require characterizing $\Theta(\delta,\mu)$ and $M(\delta,\mu)$ for non-zero levels of drift. On the contrary, such characterization is crucial for my approach and introduces two challenges.

First, drift μ affects $\Theta(\delta,\mu)$ and $M(\delta,\mu)$ by altering the stationary distribution of price gaps and changing the optimal policy (which also feeds into the gap distribution via boundary conditions). Thus, understanding the effects of drift on aggregate dynamics requires a characterization of its effects on the optimal policy $\left\{\frac{\partial \underline{z}(0)}{\partial\mu}, \frac{\partial \hat{z}(0)}{\partial\mu}, \frac{\partial \overline{z}(0)}{\partial\mu}\right\}$ and stationary density $\left(\frac{\partial f(z,0)}{\partial\mu}\right)$.

Second, as I show later, non-zero drift introduces an additional term into the expression for the cumulative output response $M(\delta,\mu)$, which is not captured when tracking firms until the first time of adjustment. I provide a way of computing this new term and generalize the approach of characterizing cumulative impulse responses, introduced by Alvarez et al. (2016), to economies with drift and asymmetries.

1.2.5 Optimal Policy of Firms under Non-Zero Drift

To approximate the optimal policy for the case of non-zero drift, I apply implicit function theorem to the system of equations that characterize the solution of the firm's problem, as discussed in Section 1.2.1. Proposition 1 states the result and the proof is provided in Appendix B.2.

Proposition 1. *Let $\sigma, \rho, \kappa > 0$. Then:*

$$\left.\frac{\partial \underline{z}(\mu)}{\partial\mu}\right|_{\mu=0} = \left.\frac{\partial \overline{z}(\mu)}{\partial\mu}\right|_{\mu=0} > 0$$
$$\left.\frac{\partial \hat{z}(\mu)}{\partial\mu}\right|_{\mu=0} > \left.\frac{\partial \overline{z}(\mu)}{\partial\mu}\right|_{\mu=0}$$

The first line states that boundaries of the inaction region move in parallel to the right as trend inflation increases. This implies that the width of the inaction region $\big(\overline{z}(\mu) - \underline{z}(\mu)\big)$ is insensitive to trend inflation at $\mu = 0$. The second line states that the return point moves in the same direction, but stronger than the boundaries. Both effects are due to the desire of the firm to stay close to the profit maximizing zero price gap for as long as possible. With a positive drift in the optimal price, the price gaps are expected to fall over time. Therefore, firms move the return point to the right to increase the total time spent in the vicinity of zero. Firms also tolerate larger positive gaps and delay adjusting because the gaps are expected to fall on their own. For the same reason, firms adjust 'sooner' for negative price gaps, as these are not expected to rise over time. Note that by taking the limit as $\rho \to 0$, I recover expressions for the no-discounting case considered in Alvarez et al. (2019).

The uneven shift of the return point and boundaries leads to asymmetry in individual adjustments. Denote the size of positive adjustments by $\Delta^+(\mu) := \hat{z}(\mu) - \underline{z}(\mu)$, and of negative adjustments by $\Delta^-(\mu) := \overline{z}(\mu) - \hat{z}(\mu)$. An immediate implication of Proposition 1 is that $\frac{\partial \Delta^+(0)}{\partial \mu} = -\frac{\partial \Delta^-(0)}{\partial \mu} > 0$, meaning that positive adjustments become larger as drift increases, whereas negative adjustments become smaller. Finally, defining asymmetry in individual adjustments by $A_I(\mu) = \frac{\Delta^+(\mu)}{\Delta^-(\mu)}$, obtains:

$$\frac{\partial A_I(0)}{\partial \mu} = \frac{2}{\overline{z}(0)} \frac{\partial \Delta^+(0)}{\partial \mu} > 0$$

This implies that asymmetry in individual price adjustments increases with trend inflation in the sense that positive adjustments become larger relative to negative adjustments. While the result might not appear surprising, it is not immediate. When exposed to a small positive drift, a firm may adjust its behavior entirely via the relative frequency of price increases and decreases, while keeping adjustments symmetric. Instead, because of a forward-looking behavior, it chooses to increase its positive adjustments in anticipation of price erosion due to trend inflation and decrease its negative adjustments for the same reason. The concerns of firms regarding future dynamics are key here: stronger discounting weakens the asymmetry, and it is entirely gone in a static model, which is the limit case as $\rho \to \infty$.

1.2.6 Impact Effect

Before stating the results for the impact effect, it is instructive to outline the channels through which drift influences the impact response of the price level. Let us decompose the impact effect of a positive shock δ, given in equation (1.1), using the definition of positive adjustments $\Delta^+(\mu) = \hat{z}(\mu) - \underline{z}(\mu)$ and performing variable substitution $z \to x := \underline{z}(\mu) - z$:

$$\Theta(\delta, \mu) = \underbrace{\Delta^+(\mu) F(\underline{z}(\mu) + \delta, \mu)}_{\text{Minimal adjustment}} + \underbrace{\int_0^\delta x f(\underline{z}(\mu) + \delta - x, \mu) dx}_{\text{Additional adjustment}} \tag{1.4}$$

If there is a positive shock of size δ, then a total mass $F(\underline{z}(\mu) + \delta, \mu)$ of agents adjust immediately, with each of them adjusting by $\Delta^+(\mu)$ at least. This is reflected in the first term and denoted by 'minimal adjustment'. Because agents are shifted strictly outside of the inaction region, their actual adjustment is larger. This 'additional adjustment' component depends on the position of the agent prior to the shock, is denoted by x and is captured by the second term. Differentiating $\Theta(\delta, \mu)$ and evaluating at $\mu = 0$ provides the following expression:

$$\frac{\partial \Theta(\delta, 0)}{\partial \mu} = \overbrace{\frac{\partial \Delta^{+}(0)}{\partial \mu} F(\underline{z}(0)+\delta, 0)}^{\text{Intensive margin}} \\ + \underbrace{\Delta^{+}(0) \frac{dF(\underline{z}(0)+\delta, 0)}{d\mu} + \int_{0}^{\delta} x \frac{df(\underline{z}(0)+\delta - x, 0)}{d\mu} dz}_{\text{Extensive margin}}$$

The effect of trend inflation on the immediate price response can be decomposed into two terms. The first is the effect on the minimal adjustment size, labeled 'intensive margin' and driven purely by changes in optimal policy. The second is the effect on the mass of adjusting agents, labeled as 'extensive margin' and driven by changes in the stationary distribution. Note that stationary density depends on optimal policy, and thus the latter will indirectly affect the extensive margin as well.[7] I provide an expression for $\frac{\partial \Theta(\delta, 0)}{\partial \mu}$ in Appendix A.1.5 and Proposition 2 states that this derivative is always positive.

Proposition 2. *Let $\sigma, \rho, \kappa > 0$. Then for any $\delta \neq 0$:*

$$\left. \frac{\partial \Theta(\delta, \mu)}{\partial \mu} \right|_{\mu=0} > 0$$

The result implies that a small positive trend amplifies the responses to positive shocks and mitigates the responses to negative shocks. Importantly, trend has a first-order effect on $\Theta(\delta, \mu)$, irrespective of shock size. Define asymmetry for impact effect analogously to individual adjustments: $A_{\Theta} = \frac{\Theta(\delta, \mu)}{-\Theta(-\delta, \mu)}$. It follows that:

$$\frac{\partial A_{\Theta}(\delta, 0)}{\partial \mu} = \frac{2}{\Theta(\delta, 0)} \frac{\partial \Theta(\delta, 0)}{\partial \mu} > 0$$

Therefore, asymmetry in the impact price responses goes up as trend inflation rises, in the sense that the magnitude of responses to positive shocks increases relative to the magnitude of responses to negative shocks.

[7] Derivatives of F and f are total (not partial) since both of their arguments depend on μ in (1.4).

Interestingly, the effect of drift on asymmetry in aggregate price responses does not vanish as shock size goes to zero:

$$\lim_{\delta \to 0} \frac{\partial A_\Theta(\delta, 0)}{\partial \mu} = \frac{2\bar{z}_0}{\sigma^2}$$

This is because the impact effect and its derivative with respect to trend inflation are of the same order for small shocks.[8] Combining a first-order approximation with respect to drift μ with a second-order approximation with respect to shock δ gives:

$$\Theta(\delta, \mu) \approx \begin{cases} \left(1 + \frac{\bar{z}_0}{\sigma^2}\mu\right)\Theta(\delta, 0) & \text{for } \delta > 0 \\ \left(1 - \frac{\bar{z}_0}{\sigma^2}\mu\right)\Theta(\delta, 0) & \text{for } \delta < 0 \end{cases}$$

This shows that drift has a multiplicative effect on the impact response. For a small positive drift, the impact effect of a positive shock is increased by $100 \cdot \frac{\bar{z}_0}{\sigma^2}\mu$ percent, whereas the response to a negative shock is decreased in the same proportion. Therefore, if a shock is small, ignoring the effect of drift produces an error of the same order as simply setting the impact response to zero.

One can also compare asymmetry at individual and aggregate levels. At the micro-level, firms react to shocks as soon as the inaction region boundaries are reached, therefore, it would be fair to compare the trend effect on individual asymmetry ($A_I(\mu)$, introduced earlier) with the trend effect on aggregate asymmetry $\left(\frac{\partial A_\Theta(\delta,0)}{\partial \mu}\right)$ for an aggregate shock δ approaching zero. The comparison yields:

$$\lim_{\delta \to 0} \frac{\partial A_\Theta(\delta, 0)}{\partial \mu} > \frac{\partial A_I(0)}{\partial \mu}$$

This follows because the trend effect on aggregate asymmetry consists of extensive and intensive margins, whereas individual asymmetry is only driven by the latter. These work in the same direction, amplifying asymmetry at the aggregate level even further.

1.2.7 Cumulative Impulse Response

An extremely useful result for cumulative impulse responses in driftless economies is that one only has to keep track of price gaps until the first adjustment.

[8] Alvarez and Neumeyer (2019) also mention that trend affects the coefficient in front of δ^2 in a second-order approximation of $\Theta(\delta, \mu)$ with respect to δ. Here, I provide an explicit expression for this interaction term for small values of μ.

Unfortunately, this result does not hold in economies with non-zero drift. Explicitly writing an infinite-horizon CIR as a limit of a finite-horizon CIR reveals that cumulative responses until finite horizon t have an additional 'tail' term, which represents cumulative deviations between the time of the last adjustment and period t. This tail term does not vanish in the limit and is not equal to zero in expectation. In the following, I derive an extension of the CIR formula for economies with non-zero drift in a more general setting, which might be useful for purposes beyond the scope of this paper.

Following Alvarez and Lippi (2019), I let $z(t)$ be an individual process on $Z = [\underline{z}, \overline{z}]$, endowed with the strong Markov property and a stationary distribution $F(z)$. I denote by $g : Z \to \mathbb{R}$ a bounded, Borel-measurable function of interest. Suppose the economy is in a steady state. In period $t = 0$, an aggregate shock distorts the distribution of z, such that distribution in $t = 0$ is given by $F_0(z)$. One can express the impulse response t periods after as follows:

$$IRF(t, F_0) = \int_{\underline{z}}^{\overline{z}} \mathbb{E}\Big(g(z(t)) - \bar{g} \mid z(0) = z\Big) dF_0(z) \quad \text{where} \quad \bar{g} = \int_{\underline{z}}^{\overline{z}} g(z) dF(z)$$

Therefore, $IRF(t, F_0)$ is the period t economy-wide average deviation of g from its steady state average $\bar{g}$ if z was initially distributed according to $F_0(z)$. Denote by $CIRF(t, F_0)$ the cumulative impulse response up to period t:

$$CIRF(t, F_0) = \int_0^t IRF(s, F_0) ds$$

Switching the order of the integration and taking the expectation operator out of the inner integral yields:

$$CIRF(t, F_0) = \int_{\underline{z}}^{\overline{z}} \mathbb{E}\Big(\int_0^t (g(z(s)) - \bar{g}) ds \mid z(0) = z\Big) dF_0(z)$$

One can first compute expected cumulative deviation of g from its steady state until t for each starting value z and then average across all starting values using the initial distribution function F_0. The statistic of interest is the infinite-horizon cumulative IRF:

$$CIRF(F_0) = \lim_{t\to\infty} CIRF(t, F_0) = \int_{\underline{z}}^{\overline{z}} \mathbb{E}\Big(\int_0^\infty (g(z(s)) - \bar{g}) ds \mid z(0) = z\Big) dF_0(z)$$

Cumulative Impulse Responses in Impulse Control Models

Now consider a special case for the process $z(t)$, namely the one resulting from an impulse control problem with a fixed return point $\hat{z}$, as in the model considered in this paper. The next proposition characterizes $CIRF(F_0)$ as a sum of two terms: the familiar expected deviation until the first adjustment and the new tail term.

Proposition 3. *Denote by $m(z)$ the expected cumulative deviation of g from its steady state $\bar{g}$ until the time of the first adjustment τ, conditional on the initial value $z(0) = z$:*

$$m(z) = \mathbb{E}\left(\int_0^{\tau} (g(z(s)) - \bar{g}) ds \;\middle|\; z(0) = z \right)$$

Let $n(t)$ be the number of adjustments between time 0 and t, so that $\tau_{n(t)}$ denotes the time of the last adjustment before t. Then:

$$CIRF(F_0) = \int_{\underline{z}}^{\overline{z}} m(z) dF_0(z) + \lim_{t\to\infty} \mathbb{E}\left(\int_{\tau_{n(t)}}^{t} (g(z(s)) - \bar{g}) ds \right)$$

Function $m(z)$ is similar to the one used to compute the cumulative IRF under zero drift. This function provides the expected cumulative deviation of g from its steady state $\bar{g}$ until the time of the first adjustment and can typically be defined with an ordinary differential equation.

To understand the new term, consider the cumulative response until some large finite time t. Each agent will have a certain number of adjustments, made until that period, denoted by $n(t)$, and the cumulative response can be split into periods before the first adjustment, in between adjustments and after the last adjustment:

$$\begin{aligned} \int_0^t (g(z(s)) - \bar{g}) ds = & \int_0^{\tau_1} (g(z(s)) - \bar{g}) ds + \\ & \sum_{i=2}^{n(t)} \int_{\tau_{i-1}}^{\tau_i} (g(z(s)) - \bar{g}) ds + \int_{\tau_{n(t)}}^{t} (g(z(s)) - \bar{g}) ds \end{aligned}$$

The idea of the proof is that when taking expectations and letting $t \to \infty$, the first term becomes the function $m(z)$, terms in the middle vanish as shown by Baley and Blanco (2020), and the last term converges to some number, which is not necessarily zero. Deviations sum up to zero in expectation if they are considered strictly in between adjustment times, as in the middle terms. This does not apply to the last term, which cumulates deviations between an adjustment time and some arbitrary t, given that the next adjustment occurs after t.

In fact, for the model considered in this paper and $g(z) = z$, this tail term is equal to zero if and only if $\mu = 0$. This is because if $\mu \neq 0$, then the return point $\hat{z}$ is not equal to the average gap $\bar{x} = \int_{\underline{z}}^{\overline{z}} z dF(z)$, which implies that the expected cumulative deviation between the last time of adjustment $\tau_{n(t)}$ and arbitrary t is not equal to zero. For example, if t is very close to $\tau_{n(t)}$, then for any time $s \in (\tau_{n(t)}, t)$, $z(s)$ is very close to the return point $\hat{z}$ in expectation, and thus relatively far from the average gap $\bar{x}$, so that the expected cumulative deviation $\mathbb{E}\big(\int_{\tau_{n(t)}}^{t}(z(s) - \bar{x})ds\big)$ is non-zero.

Computing the Tail Term

Unlike the cumulative response until the first adjustment, the tail term does not allow for an immediate characterization. However, there are at least two ways of dealing with this issue. The first one relies on the fact that the new term does not depend on the initial distribution F_0, as it considers paths after the first adjustments. These paths are independent of the initial condition due to the strong Markov property of $z(t)$. Note that setting the initial distribution F_0 equal to the stationary distribution F results in zero impulse response by definition:

$$CIRF(F) = \int_{\underline{z}}^{\overline{z}} m(z)dF(z) + \lim_{t\to\infty} \mathbb{E}\left(\int_{\tau_{n(t)}}^{t} (g(z(s)) - \bar{g})ds\right) = 0$$

This allows to obtain an expression for the limit term and express the cumulative response as follows:

Corollary 1.

$$CIRF(F_0) = \int_{\underline{z}}^{\overline{z}} m(z)dF_0(z) - \int_{\underline{z}}^{\overline{z}} m(z)dF(z)$$

When there is no drift, the inaction region is symmetric ($\underline{z} = -\overline{z}$), $F(z)$ is a symmetric distribution, and $m(z)$ exhibits negative symmetry ($m(-z) = -m(z)$). These imply that $\int_{\underline{z}}^{\overline{z}} m(z)dF(z) = 0$ and $CIRF(F_0) = \int_{\underline{z}}^{\overline{z}} m(z)dF_0(z)$. Therefore, one only needs to track each agent until the first time of adjustment to compute the entire cumulative response. However, if drift is non-zero, one can still consider paths until the first adjustment, but must additionally subtract the average cumulated deviation under the stationary distribution.

The second approach uses the notion of a discounted cumulative impulse response:

$$DCIRF(r, F_0) = \lim_{t\to\infty} DCIRF(r, t, F_0) = \int_0^{\infty} e^{-rs} IRF(s, F_0)ds$$

which allows to eliminate the tail term for any $r > 0$. Then, the usual CIR can be expressed as a limit case of its discounted counterpart:

$$CIRF(F_0) = \lim_{r \to 0} DCIRF(r, F_0)$$

This approach substitutes the inconvenient tail term with a more tractable one, and the next proposition provides the result.

Proposition 4. *Denote by $m(r,z)$ the expected discounted cumulative deviation of g from its steady state until the time of the first adjustment τ, conditional on initial value $z(0) = z$:*

$$m(r,z) = \mathbb{E}\left(\int_0^\tau e^{-rs}(g(z(s)) - \bar{g})ds \;\middle|\; z(0) = z \right)$$

Then:

$$CIRF(F_0) = \int_{\underline{z}}^{\overline{z}} m(0,z)dF_0(z) + \frac{1}{\mathbb{E}\left(\tau \mid z(0) = \hat{z}\right)} \lim_{r \to 0} \frac{m(r,\hat{z})}{r}$$

Here, the first term is the same as before, since trivially $m(0,z) = m(z)$, whereas the second term provides an alternative way of computing the tail term from Proposition 3.

The first approach from Corollary 1 uses all familiar objects but requires computing the integral of $m(z)$ twice – under initial and stationary distributions. Using the second approach from Proposition 4, one needs to compute an additional function $m(r,z)$, which can typically be defined with an ordinary differential equation, similar to $m(z)$. Therefore, each approach may be more or less preferable, depending on the application. For example, if one deals with shocks that shift the stationary distribution (as in this paper), then the first approach provides a much easier way of computing CIR because F_0 inherits the shape of F. On the other hand, if the initial distribution is not related to the stationary one, it might be more convenient to analyze CIR using the second approach, as it only requires computing the integral under F_0 (although F is still required to compute the steady state average $\bar{g}$).

To determine whether the tail term is qualitatively important and whether one can omit it for simplicity, recall that it neither depends on the initial distribution F_0 nor interacts with it. This implies that it corrects for the level of the cumulative response, acting as an intercept. Therefore, omitting it not only changes the CIR value, but may also flip its sign if the true value is sufficiently close to zero. In the next section, I discuss the special importance of the tail term for cumulative responses to δ shocks considered in this paper.

Application to δ Shocks

A δ shock considered in this paper shifts the stationary distribution F in parallel. The initial distribution F_0 is given by the stationary distribution F, shifted by δ and truncated to the inaction region, together with a mass point at $\hat{z}$, which is due to firms that adjust on impact. As noted earlier, it is most convenient in this situation to use Corollary 1 for computing the cumulative response, so that for $\delta > 0$ it is given by:

$$CIRF(\delta) = \int_{\underline{z}}^{\overline{z}-\delta} m(z)dF(z+\delta) + \underbrace{m(\hat{z})F(\underline{z}+\delta)}_{=0} - \int_{\underline{z}}^{\overline{z}} m(z)dF(z) \tag{1.5}$$

where $m(\hat{z}) = 0$, as shown in Baley and Blanco (2020). Note that the tail term does not affect the *slope* of CIRF(δ), which is entirely determined by cumulative deviations until the first adjustment, captured in the first term. Instead, it shifts the entire function, which has special importance for very small and very large shocks. If $\delta = 0$, then ignoring the tail term would give that $CIRF(0) = \int_{\underline{z}}^{\overline{z}} m(z)dF(z)$, which may not be equal to zero, implying a 'response' despite the absence of a shock. If δ is large, so that $\delta \geq (\overline{z} - \underline{z})$, then the first term in (1.5) vanishes, and the CIR is entirely determined by the tail term. Omitting the tail term would imply that $CIRF(\delta) = 0$ for all $\delta \geq (\overline{z} - \underline{z})$, whereas it might be different from zero.

1.2.8 Sensitivity of the Cumulative Impulse Response to Drift

I now use results from the previous section to study the sensitivity of the cumulative output response to drift μ. Recall from section 1.2.2 that it is given by:

$$M(\delta, \mu) = -\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \mathbb{E}\left(\int_0^\infty (z(t) - \bar{x}(\mu))\, dt \;\middle|\; z(0) = z\right) dF_\delta(z, \mu)$$

where $\bar{x}(\mu) = \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} z dF(z, \mu)$. Using Corollary 1 and writing $M(\delta, \mu)$ as in (1.5) yields:

$$M(\delta, \mu) = \int_{\underline{z}(\mu)}^{\overline{z}(\mu)-\delta} m(z, \mu) f(z+\delta, \mu) dz - \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} m(z, \mu) f(z, \mu) dz$$

where

$$m(z, \mu) = -\mathbb{E}\left(\int_0^\tau (z(t) - \bar{x}(\mu))dt \;\middle|\; z(0) = z\right)$$

Function $m(z, \mu)$ solves the following differential equation:

$$z - \bar{x}(\mu) = -\mu m_z(z, \mu) + \frac{\sigma^2}{2} m_{zz}(z, \mu)$$

with boundary conditions $m(\underline{z}(\mu), \mu) = m(\overline{z}(\mu), \mu) = 0$. Proposition 5 states that drift has a first-order effect on the cumulative output response, irrespective of the shock size.

Proposition 5. *Let $\sigma, \rho, \kappa > 0$. Then for any $\delta \neq 0$:*

$$\left.\frac{\partial M(\delta, \mu)}{\partial \mu}\right|_{\mu=0} < 0$$

In accordance with the results on impact effect $\Theta(\delta, \mu)$, trend inflation amplifies price responses to positive shocks and thus mitigates responses of output, which is reflected by the negative sign of the derivative. The reverse is true for negative shocks, as in this case output responses are amplified. Note that the tail term is crucial for this result, without it, the derivative is positive for small shocks, zero for large shocks and negative for intermediate values.

Define asymmetry in CIR as a difference in magnitudes of responses to positive and negative shocks: $A_M(\delta, \mu) = M(\delta, \mu) - (-M(-\delta, \mu))$. Here, I am using difference instead of ratio in order to ensure that asymmetry is well-defined for shocks of any size because $M(\delta, 0) = 0$ for all δ such that $|\delta| > 2\overline{z}_0$. It is, however, also possible to define it as a ratio, provided $M(\delta, \mu) > 0$ and $M(-\delta, \mu) < 0$. It follows immediately that:

$$\frac{\partial A_M(\delta, 0)}{\partial \mu} = 2\frac{\partial M(\delta, 0)}{\partial \mu} < 0$$

Therefore, cumulative output responses to positive shocks become smaller relative to cumulative output responses to negative shocks as trend inflation increases.

To determine whether the drift effect is sizable, I combine a first-order approximation of $M(\delta, \mu)$ with respect to μ and a second-order approximation with respect to δ:

$$M(\delta, \mu) \approx \begin{cases} \left(1 - \frac{|\delta|}{\sigma^2}\mu\right) M(\delta, 0) & \text{for } \delta > 0 \\ \left(1 + \frac{|\delta|}{\sigma^2}\mu\right) M(\delta, 0) & \text{for } \delta < 0 \end{cases}$$

Cumulative output response is amplified by $100 \cdot \frac{|\delta|}{\sigma^2}\mu$ percent if a shock is negative and mitigated in the same proportion for a positive shock. Thus, for small shocks, the drift effect is negligible, but it becomes more important as the shock size increases. For large shocks, the drift might not only amplify or mitigate output responses, but also change their sign. This result is discussed and formalized in the following section.

1.2.9 Large Shocks

The drift effect is of particular importance for large shocks. As noted previously, in the driftless case, the price level reacts one-to-one to a large nominal shock on impact, which results in monetary neutrality, i.e., output is not affected by the shock. These results break down when trend inflation is non-zero.

To see why this occurs, consider a positive shock δ that is large in the sense that it shifts the entire distribution outside of the inaction region ($\delta \geq \overline{z}(\mu) - \underline{z}(\mu)$). The impact effect for this shock is given by:

$$\Theta(\delta, \mu) = \delta + \hat{z}(\mu) - \bar{x}(\mu) \quad \text{where} \;\; \bar{x}(\mu) = \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} z f(z, \mu) dz$$

The entire distribution of price gaps is initially shifted to the left by δ, and the mean price gap immediately after the shock and before the adjustment becomes $\bar{x}(\mu) - \delta$. Because all firms are pushed outside the inaction region, the aggregate adjustment equals $\hat{z}(\mu) - (\bar{x}(\mu) - \delta)$, which gives the impact effect. When $\mu = 0$, both the average gap $\bar{x}(0)$ and the return point $\hat{z}(0)$ are zero, and thus the impact effect is equal to the shock, meaning that the aggregate price responds one-to-one. If trend inflation is positive ($\mu > 0$), then $\hat{z}(\mu) > 0$ and $\bar{x}(\mu) < 0$, which leads to price overreaction: $\Theta(\delta, \mu) > \delta$. Note that if the shock is negative and $\delta \leq -|\overline{z}(\mu) - \underline{z}(\mu)|$, then $|\Theta(\delta, \mu)| < |\delta|$ and there is no overreaction. All of the above has implications for the cumulative output response $M(\delta, \mu)$. Proposition 6 formalizes the results.

Proposition 6. *Let $\sigma, \rho, \kappa > 0, \mu > 0$ and sufficiently small.*

Then there exist $\delta_\Theta(\mu), \delta_M(\mu) > 0$ such that $\delta_\Theta(\mu), \delta_M(\mu) < \overline{z}(\mu) - \underline{z}(\mu)$ and:

$$\begin{aligned} \Theta(\delta, \mu) > \delta \quad & \textit{for all} \;\; \delta > \delta_\Theta(\mu) \\ M(\delta, \mu) < 0 \quad & \textit{for all} \;\; \delta > \delta_M(\mu) \end{aligned}$$

Proposition 6 states that if trend inflation is positive and small, then there exist thresholds $\delta_\Theta(\mu)$ and $\delta_M(\mu)$ such that price overreacts on impact to any positive shock larger than $\delta_\Theta(\mu)$ and the cumulative output response is negative for any positive shock larger than $\delta_M(\mu)$.[9] The latter implies that positive shocks eventually become contractionary if trend inflation is positive. Crucially, the thresholds are both strictly smaller than the width of the inaction region $\overline{z}(\mu) - \underline{z}(\mu)$. Therefore, a positive shock does not have to shift the entire distribution outside of the inaction region to induce price overshooting and a decline in output when trend inflation is positive, as these effects are already achieved for smaller shocks.

1.2.10 Summary and Comparison with the Driftless Case

In the following I summarize and illustrate the main analytical results of the paper. Figure 1.2.3 plots the impact price response $\Theta(\delta, \mu)$ (left panel) and the cumulative response of output $M(\delta, \mu)$ (right panel) against different values of the δ shock, normalized by the width of the inaction region.[10] I also normalize the impact effect to maintain comparability with the size of the shock. The solid blue lines correspond to the driftless case $\mu = 0$, and the red dashed lines correspond to $\mu = 0.1$. Recall the three properties of the impact effect and the cumulative output response under zero trend inflation, discussed in section 1.2.3: (1) both statistics are symmetric for positive and negative shocks, (2) the size of the impact effect $\Theta(\delta, 0)$ is always weakly smaller than the shock size, and cumulative output response $M(\delta, 0)$ to positive shocks is always weakly positive, and (3) if $|\delta|$ is larger than the width of inaction region $(\overline{z} - \underline{z})$, then $\Theta(\delta, 0) = \delta$ and $M(\delta, 0) = 0$. All these properties are illustrated in Figure 1.2.3 by the solid blue lines.

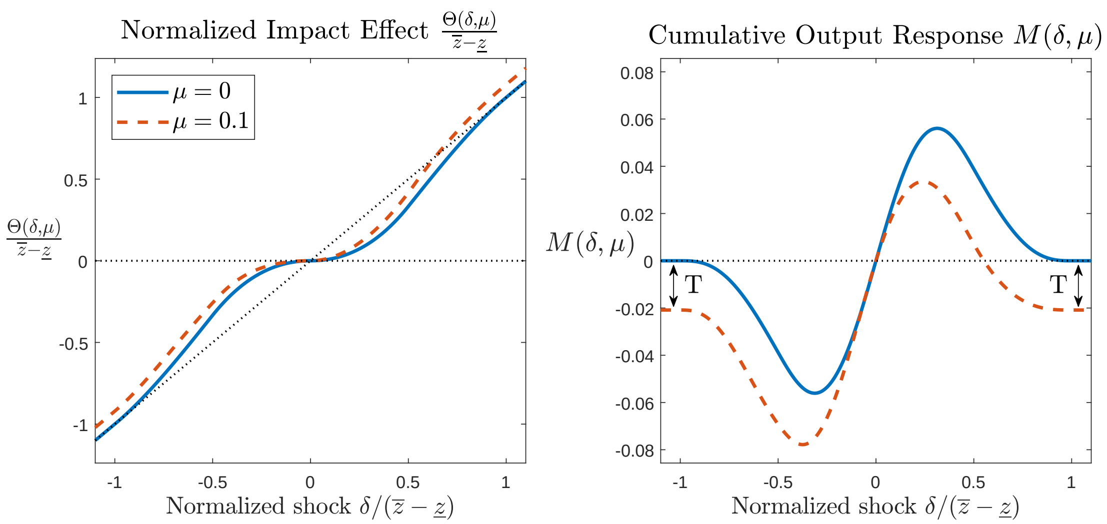
Now consider the case of positive trend inflation. The responses to positive and negative shocks are asymmetric, so that the first property does not hold anymore. The impact price responses to positive shocks are amplified, whereas the responses to negative shocks are mitigated compared to the driftless benchmark. This can be seen on the left panel, as the dashed red line lies above the solid blue one. The opposite is true for the cumulative output responses, which become stronger after negative shocks and weaker after positive ones as trend inflation rises. This is depicted on the right panel, where the dashed red line lies below the solid blue one.

Furthermore, large positive shocks lead to price overreaction on impact and cause negative cumulative output responses, which invalidates both the

[9] Alvarez and Neumeyer (2019) show that price response exceeds the shock on impact if $\mu \to \infty$ and provide numerical examples when this happens for finite values of μ.

[10] Both statistics are exact values and not first-order approximations with respect to drift μ.

Figure 1.2.3: Impact and Cumulative Responses



Left panel: impact price response $\Theta(\delta, \mu)$, right panel: cumulative output response $M(\delta, \mu)$. X-axis: shocks δ normalized by the width of the inaction region $\overline{z} - \underline{z}$. Impact effect $\Theta(\delta, \mu)$ is also normalized by $\overline{z} - \underline{z}$ for comparability of x- and y-axes. Solid blue lines: $\mu = 0$, dashed red lines: $\mu = 0.1$. The tail term is denoted by T on the right panel. Rest parameter values: $\sigma^2 = 0.05$, $\rho = 0.05$, $\kappa = 0.05$. Threshold values: $\delta_\Theta(0.1) = 0.64(\overline{z} - \underline{z})$ and $\delta_M(0.1) = 0.55(\overline{z} - \underline{z})$.

second and third properties. Proposition 6 states that there are two thresholds, $\delta_\Theta(\mu)$ and $\delta_M(\mu)$, such that shocks larger than these thresholds cause price overshooting and output contraction, respectively. The former is determined by the intersection of the red dashed line and a 45° line on the left panel of Figure 1.2.3, where $\Theta(\delta, \mu) = \delta$. The latter threshold corresponds to the point where the dashed red line crosses zero on the right panel of Figure 1.2.3 for a positive shock δ, so that $\delta > 0$ and $M(\delta, \mu) = 0$. Numerical computation yields that $\delta_\Theta(0.1) = 0.64(\overline{z} - \underline{z})$ and $\delta_M(0.1) = 0.55(\overline{z} - \underline{z})$. Therefore, when $\mu = 0.1$, any shock δ larger than 64% of the width of the inaction region causes price overshooting on impact and any shock larger than 55% of the width of the inaction region leads to a cumulative contraction in output. The higher the trend inflation, the harder it is to stimulate output, as even medium-size positive nominal shocks have a reversed effect. The threshold for the reversed effect on output is smaller than the threshold for price overshooting. Thus, there is a range of shocks (55% - 64% of the width of the inaction region) for which cumulative output response is negative, even though price level responds less than one-to-one on impact. On the contrary, negative shocks never lead to an expansion in output and price overshooting. The price level always underreacts to negative shocks on impact and the cumulative output response is always negative if trend inflation is positive.

I lastly note the role of the tail term that appears in the expression for the cumulative output response when trend inflation is non-zero. As discussed in section 1.2.7, the CIR is entirely determined by this term if $|\delta| \geq (\overline{z} - \underline{z})$. Therefore, one can directly see the tail term in Figure 1.2.3, where it is denoted by T on the right panel. Not only is it quantitatively important, but it is also the only source of difference between the cases of positive and zero trend inflation.

The results of this section are derived under the assumption that firms can only adjust prices at a fixed menu cost. This assumption is not pivotal for the results, although it substantially simplifies the analysis and exposition. In Appendix A.1.8 I extend the model to allow for random opportunities of costless adjustments and show that some of the results (including the overshooting result of Proposition 6) can also be proven in this richer setting. Such an extension is usually referred to as "CalvoPlus" in the literature, as it nests both the traditional Calvo (1983) model and the standard menu cost model. An even broader class of models with generalized hazard functions is studied in Caballero and Engel (2007) and Alvarez et al. (2020), and extending the results to their setting remains a challenge.

1.2.11 Shock to the Drift

The analytic results of this paper allow one to study the dynamics of the price level after a different type of shock, namely a shock to the trend inflation. Consider a driftless economy in a steady state. Imagine that at date $t = 0$ the trend inflation is unexpectedly changed to some small positive value, so that $\mu > 0$ from now on. How would the economy respond to such a change on impact and what would be its cumulative effect?

First, the increase in drift induces a change in firms' policy. The inaction region shifts to the right and therefore a mass of firms adjusts prices upward immediately. Formally, this is captured by the impact effect $\Theta(\mu)$ which now only depends on the change in the drift:

$$\Theta(\mu) = \int_{\underline{z}(0)}^{\underline{z}(\mu)} \big(\hat{z}(\mu) - z\big) f(z,0) dz$$

Since $\hat{z}(\mu) > \underline{z}(\mu) > \underline{z}(0)$, the impact effect of an increase in drift on the price level is positive. Even though the desired price p^* is unchanged at

$t = 0$, the aggregate price goes up in anticipation of future rise in p^* due to positive drift. However, the impact response is only of second order in μ, since $\frac{\partial \Theta(\mu)}{\partial \mu}\big|_{\mu=0} = 0$, which can be easily verified.

Computing cumulative impulse responses is less straightforward. Unlike in the case of a nominal δ shock, a shock to the drift changes the steady state of the economy, so that cumulative impulse response of output diverges. One can nevertheless compute the CIR of the price level, considering its deviations from the new long-run trend. Although this statistic does not have a clear economic interpretation, it is still informative about the speed at which the aggregate price converges to the new trend. Such CIR is given by:

$$M(\mu) = \int_{\underline{z}(\mu)}^{\overline{z}(0)} m(z,\mu) f(z,0) dz - \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} m(z,\mu) f(z,\mu) dz$$

where

$$m(z,\mu) = \mathbb{E}\left(\int_0^{\tau} (z(t) - \bar{x}(\mu)) dt \;\middle|\; z(0) = z \right)$$

In Appendix A.1.7 I show that $\frac{\partial M(\mu)}{\partial \mu}\big|_{\mu=0} > 0$. It implies that after a small increase in drift μ, the aggregate price level stays above its new trend for some time, despite the initial jump $\Theta(\mu)$ being negligible.

A potential way to study the dynamics of output is to consider discounted cumulative responses, as these are finite even in the presence of shifts in the steady state levels. Given the initial increase in the price level, the short-run effect on output is negative, whereas long-run effects may be positive. I leave the computations of the discounted CIR for future research.

1.3 Empirical Evidence

In this section I test several predictions of the model derived above. I start with the effect of trend inflation on aggregate responses to shocks, as these are the main focus of the paper. However, a substantial part of the mechanism operates via changes in firm behavior induced by the presence of drift. Therefore, I also provide evidence for the relationship between drift and asymmetry in micro-level adjustments. As I show, many of the micro- and macro-level implications of the theory are supported by the data.

1.3.1 Drift and Asymmetry at the Macro Level

I start by testing whether trend inflation affects asymmetry in aggregate responses to monetary shocks. I use monthly sectoral data on Producer

Price Index (PPI) provided by the Bureau of Labor Statistics, as well as data on Industrial Production (IP) provided by the Federal Reserve System. To estimate impulse responses I use local projections as in Jordà (2005). This approach has been widely utilized in the literature to test for asymmetries, non-linearities and state-dependence of impulse responses (Auerbach and Gorodnichenko (2012), Ramey and Zubairy (2014), Tenreyro and Thwaites (2016)). The main advantage of local projections is the ease of inclusion of non-linear terms, which are of central interest in this paper. The baseline shock measure is the one computed by Jarociński and Karadi (2020) using high frequency identification and separating monetary policy shocks from central bank information shocks. In Appendix A.3.4 I show that results are generally robust to alternative shock measures.

The central idea is to exploit cross-sectoral heterogeneity in trend inflation to see whether it relates to asymmetry in production and price responses. To ensure that impulse responses for every subset of industries are estimated on the same set of shocks, I use a balanced panel. The sample spans between February 1990 and January 2013 and consists of 52 industries.[11] I provide more details on data construction and properties of IP and PPI In Appendix A.3.1.

Asymmetric Responses

The simplest way of introducing asymmetry is estimating piecewise linear impulse responses with a kink at zero by including positive and negative shocks separately in the regression. To avoid ambiguity, I will refer to interest rate cuts (monetary easing) as 'positive' shocks, whereas to interest rate hikes (monetary tightening) as 'negative' shocks. Thus, the sign of a shock corresponds to the intended effect on output, which provides the following non-linear panel local projection:

$$y_{i,t+h} - y_{i,t-1} = \alpha_{i,h} + \beta_h^P \max(\varepsilon_t, 0) + \beta_h^N \min(\varepsilon_t, 0) + \gamma_h' \mathbf{x}_{i,t} + \nu_{i,t+h} \quad (1.6)$$

where $y_{i,t+h} - y_{i,t-1}$ is the growth rate of the dependent variable (IP or PPI) between $t-1$ and $t+h$ in industry i, $\alpha_{i,h}$ is an industry fixed effect, ε_t is the monetary policy shock, and $\mathbf{x}_{i,t}$ is a vector of controls. This specification directly estimates cumulative impulse responses. The monetary shocks are scaled and normalized such that positive values correspond to interest rate cuts and a shock of size one represents a one standard deviation shock. Here,

[11] Even though interest rates have stayed low in 2009 – 2013, this period is still informative as it features both positive and negative monetary shocks (see Figure A.3.5 in Appendix A.3.3). Considering the period until June 2008 does not alter the main results.

β_h^P provides the impulse response to a one standard deviation positive shock h periods after impact, and $(-\beta_h^N)$ is the response to a negative shock of the same size. Note that the standard theory predicts that $\beta_h^P > 0$ and $\beta_h^N > 0$. The set of controls includes a time trend, contemporaneous and lagged growth rates of aggregate industrial production and of a commodity price index, as well as lags of the monetary shock, effective federal funds rate, and industry-specific growth rates of IP and PPI. I set the lag length to 6 months and also include contemporaneous industry-specific growth rate of IP in the PPI projection and vice versa.[12] Finally, I smooth impulse responses with a 5-month centered moving average when plotting them, in order to ease comparisons.[13]

The preferred measure of asymmetry is the ratio between the magnitudes of responses to positive and negative shocks, given by β_h^P/β_h^N, because it controls for the size of an average (linear) response. Values below one indicate that positive shocks have a smaller effect relative to negative shocks, and larger deviations from one correspond to stronger degrees of asymmetry. However, this measure is only meaningful if both β_h^P and β_h^N are positive. Whenever this condition is violated, I have to use an alternative measure, defined as a difference in magnitudes $(\beta_h^P - \beta_h^N)$. In this case, negative values indicate that monetary tightening has stronger effects compared to monetary easing.

As a first step, I estimate (1.6) on the entire sample. Figure 1.3.1 plots the impulse responses of industrial production (top row) and PPI (bottom row) to one standard deviation monetary shock.[14] The dashed red lines depict responses to negative shocks $(-\beta_h^N)$, whereas the solid blue lines show the negatives of responses to positive shocks $(-\beta_h^P)$ to ease comparison. In the right column, I plot asymmetries in responses to positive and negative shocks. I employ the preferred measure of asymmetry (ratio) for industrial production, but have to use the alternative (difference) for PPI because these responses switch signs.

Industrial production exhibits a strong and significant degree of asymmetry, with negative shocks having a much larger effect on IP than positive shocks. At the horizon of 12 months, a one standard deviation negative shock has a five times stronger effect on production than a positive shock

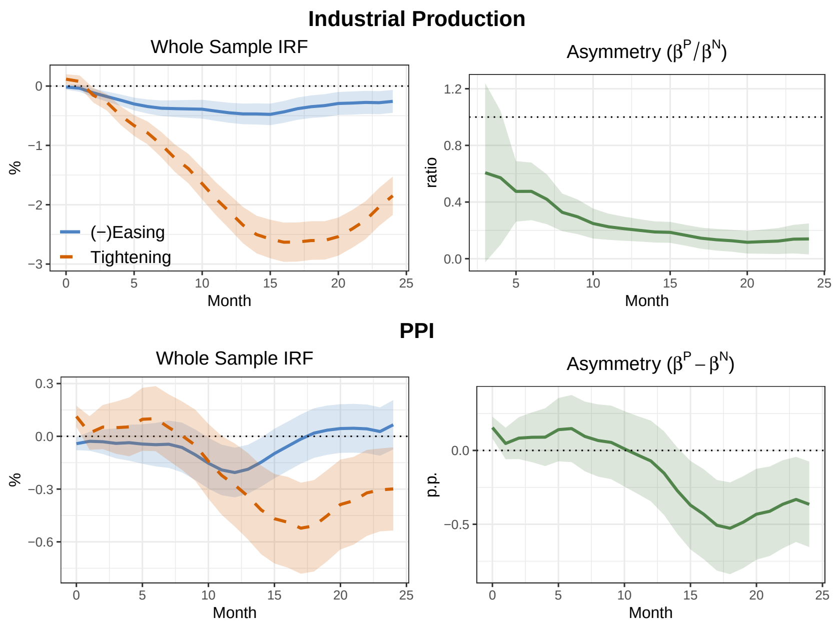
[12] The set of controls is standard, and I consider a much smaller set of controls as a robustness check in Appendix A.3.4. The commodity price index is the one produced by the Commodity Research Bureau and used in Coibion (2012) (data taken from the website of Valerie Ramey https://econweb.ucsd.edu/∼vramey/research.html#data)

[13] This is a common practice in the literature, it does not affect the results, and the unsmoothed plots are presented in Appendix A.3.4.

[14] In the sample, monetary shocks have a standard deviation of 4.8 basis points.

of the same size. There is less evidence for asymmetry in PPI responses, although at longer horizons negative shocks tend to cause larger responses than positive ones. The previous literature has focused on asymmetries at the aggregate level, and similar patterns have been documented by Angrist et al. (2018) and Tenreyro and Thwaites (2016), among others. Results in Figure 1.3.1 suggest that asymmetries found in aggregate data are also present at the sectoral level. I now turn to the interaction between trend inflation and asymmetry in responses.

Figure 1.3.1: Piecewise Linear Cumulative Impulse Responses, Entire Sample



Impulse responses of industrial production (top row) and PPI (bottom row) to one standard deviation monetary shock, estimated on the entire sample. Piecewise linear specification as in (1.6). Dashed red lines: responses to a negative shock, solid blue lines: negatives of responses to a positive shock. Right column: asymmetry in responses, measured as the ratio of magnitudes for IP (positive over negative) and as the difference in magnitudes for PPI (positive minus negative). The shaded areas correspond to 68% confidence intervals, based on Newey-West standard errors. The standard errors for asymmetry are computed by the delta method.

To determine whether asymmetry is affected by trend inflation, I compute trend inflation for each industry as an average PPI growth rate over the entire period, and split the sample into two groups: industries with trend inflation above and below the median.[15] The 'low' inflation group has an average

[15]In addition, I omit the top and bottom 2.5% of industries in terms of trend inflation

(median) trend inflation of 1.79% (1.85%) p.a., whereas for the 'high' inflation group the numbers are 3.44% and 3.22% respectively.

For the next step, I estimate (1.6) separately for each group. Figure 1.3.2 summarizes the results for industrial production (top row) and PPI (bottom row). The first column provides responses of industries with trend inflation below the median, whereas the second column shows those with trend inflation above the median. As before, I plot negatives of responses to positive shocks to ease comparison. The third column compares the asymmetry between responses to positive and negative shocks in the two groups. Again, I employ the preferred measure (ratio) for industrial production and have to use an alternative (difference) for PPI. The solid green lines correspond to the high trend inflation group, the dashed yellow lines represent the low trend inflation group, and the dotted black lines show asymmetry in the entire sample.

Firstly, PPI in the low inflation industries exhibits negative asymmetry, whereas in the high inflation industries asymmetry is predominantly positive. Sectors with low trend inflation do not raise prices after positive shocks, but decrease them substantially after negative ones. The opposite is observed for industries with high trend inflation: in the first year after impact, positive shocks have a much larger effect than negative shocks, whereas at longer horizons effects are not significantly different. These findings are in line with the theoretical predictions of this paper: higher trend inflation amplifies price responses to positive shocks and mitigates reaction to negative shocks.

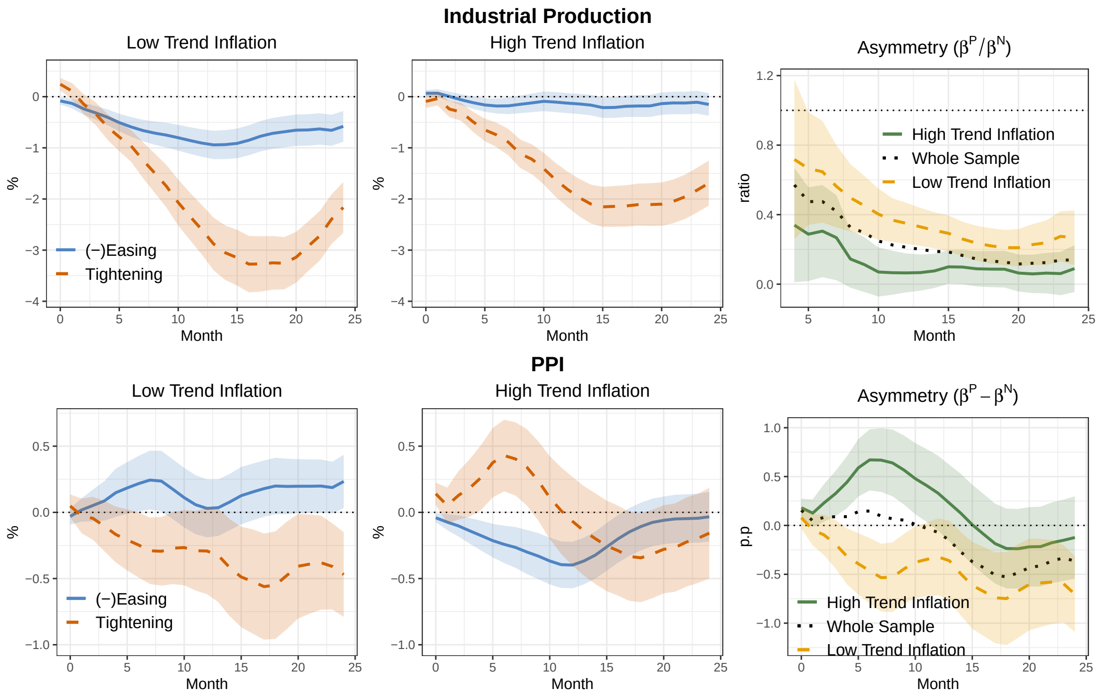
Secondly, there is a substantial difference in the asymmetry of industrial production responses between the two groups. In the low inflation sectors, positive shocks have a strong and significant effect on IP, whereas among sectors with high trend inflation their effect is more than halved and barely significant. Negative shocks also have a smaller but nevertheless pronounced and significant effect in the latter sample. This indicates that higher trend inflation is related to overall weaker effects of monetary shocks on industrial production, which is also found by Ascari and Haber (2020) in aggregate data, who use time variation in trend inflation. However, this drop in overall policy potency is disproportionately split between positive and negative shocks, as shown in the third panel, depicting asymmetry.

In both groups, the ratio between the magnitudes of responses to positive and negative shocks lies below one, but is much smaller for industries with high trend inflation. For example, compare the asymmetries at a 12-month horizon. In the low inflation group, interest rate cuts cause a three times

---

from the original sample in order to control for outliers. Results are robust to a more conservative trimming, as well as to using the entire sample (see Appendix A.3.4).

Figure 1.3.2: Piecewise Linear Cumulative Impulse Responses for Low and High Trend Inflation Industries



Impulse responses of Industrial Production (top row) and PPI (bottom row) to one standard deviation monetary shock in industries with trend inflation below the median (left column) and above the median (central column). Piecewise linear specification as in (1.6). Dashed red lines: responses to a negative shock, solid blue lines: negatives of responses to a positive shock. Third column: asymmetry in responses, measured as the ratio of magnitudes for IP (positive over negative) and as the difference in magnitudes for PPI (positive minus negative). Solid green lines: industries with trend inflation above the median, dashed yellow lines: below the median, dotted black lines: entire sample. The shaded areas correspond to 68% confidence intervals, based on Newey-West standard errors. The standard errors for asymmetry are computed by the delta method.

weaker response than interest rate hikes. In the high inflation group, the impact of positive shocks is more than 10 times weaker than the impact of negative shocks. Although the difference between the two groups is significant only at medium term horizons, there is a clear and consistent distance between the point estimates at all horizons. This is in line with the theoretical prediction of the model, which states that trend inflation reduces the relative strength of positive shocks on output.

Asymmetry and Shock Size

So far I estimated piecewise linear impulse responses, focusing on asymmetry in reactions to positive and negative shocks, irrespective of their size. The theoretical results, however, highlight the importance of non-linearities and their interactions with trend inflation. The size of a shock is especially important for output responses. While price reaction in the model is always increasing in the shock size, the output response is non-monotonic, i.e., it grows for small shocks and falls when shocks are large. In the latter case, trend inflation plays a special role, as large positive shocks may lead to contractions in output under positive trend inflation. To determine whether this holds empirically, I now add non-linear terms to local projections for industrial production.

To allow for non-monotonicity of production impulse responses for both positive and negative shocks, at least a third-order polynomial is required.[16] I estimate the following non-linear panel local projection:

$$IP_{i,t+h} - IP_{i,t-1} = \alpha_{i,h} + \beta_{1h}\varepsilon_t + \beta_{2h}\varepsilon_t^2 + \beta_{3h}\varepsilon_t^3 + \gamma_h'\mathbf{x}_{i,t} + \nu_{i,t+h} \quad (1.7)$$

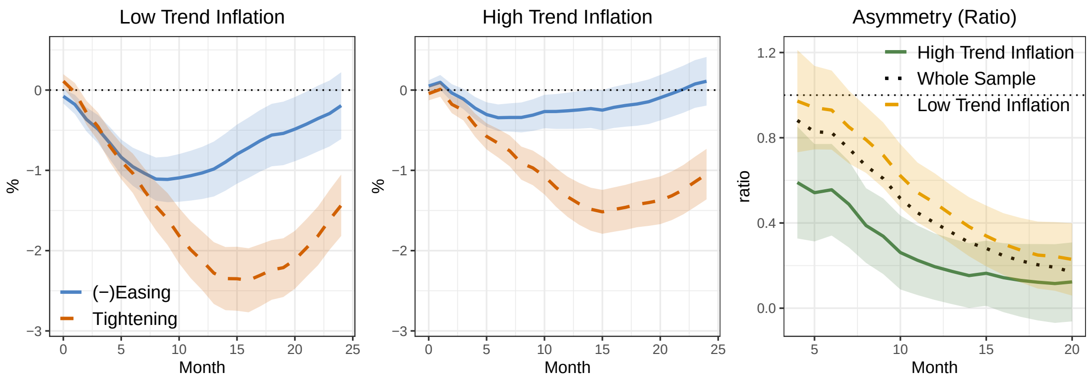
where $IP_{i,t+h} - IP_{i,t-1}$ is the growth rate of industrial production between $t-1$ and $t+h$, $\alpha_{i,h}$ is an industry fixed effect, ε_t is the monetary policy shock and $\mathbf{x}_{i,t}$ is a vector of controls, which is the same as before.

Firstly, I plot the impulse responses to one standard deviation positive and negative shocks in Figure 1.3.3 to determine whether the findings of the previous section are robust to an alternative projection specification. The results closely resemble those obtained using piecewise linear local projections, depicted in Figure 1.3.2.

Secondly, making use of non-linearity, I plot the impulse responses for 6-, 12- and 24-months horizons for different shock values in Figure 1.3.4. The x-axis corresponds to the shock values between -2 and 2 standard deviations. The y-axis shows the impulse responses as functions of the shock value. The left panel depicts the impulse responses 6 months after impact, which are close to linear for both groups, but already exhibit small asymmetry. At the 12-months horizon, asymmetry strengthens, especially for the high inflation group. In these industries, positive shocks have a very small impact on production, whereas negative shocks lead to substantial responses. Therefore, production responses to positive shocks estimated on the entire sample are almost entirely driven by industries with low trend inflation. In addition, the

[16] In addition, a second-order polynomial would always result in larger degrees of asymmetry for larger shocks. In Appendix A.3.4 I show that the results are robust to including higher order terms.

Figure 1.3.3: Non-Linear Cumulative Impulse Responses of Industrial Production



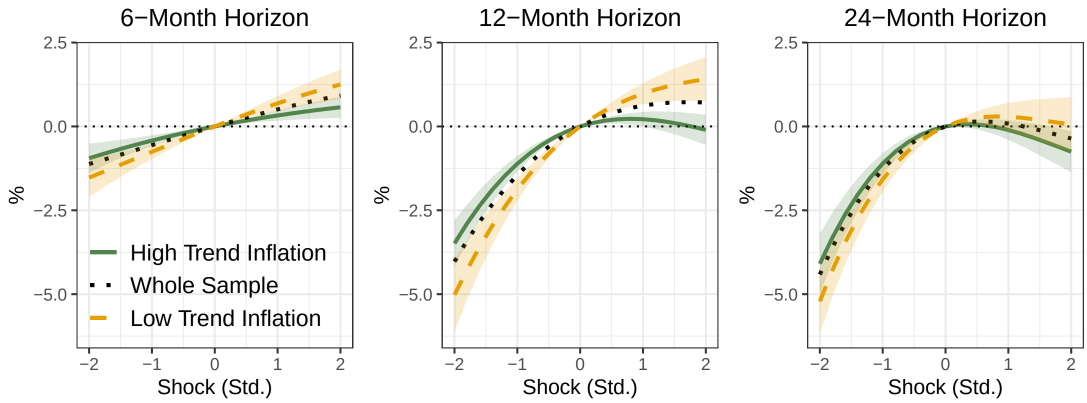
Impulse responses of industrial production to one standard deviation monetary shock in industries with trend inflation below the median (left panel) and above the median (central panel). Non-linear specification as in (1.7). Dashed red lines: responses to a negative shock, solid blue lines: negatives of responses to a positive shock. Third panel: asymmetry in responses, measured as the ratio of magnitudes (positive over negative). Solid green line: industries with trend inflation above the median, dashed yellow line: below the median, dotted black line: entire sample. The shaded areas correspond to 68% confidence intervals, based on Newey-West standard errors. The standard errors for asymmetry are computed by the delta method.

curve in the high inflation group bends toward zero as positive shocks become larger, which does not happen among sectors with low trend inflation.

At the 24-months horizon, production falls after large positive shocks in the high inflation sectors, but its response remains positive in the low inflation group. In contrast, negative shocks always lead to output contractions in both groups, even though the polynomial permits reversals for both positive and negative shocks simultaneously. This shape of the impulse response curve persists as I increase polynomial order, allowing for more flexibility, and is a robust feature of the data.

Altogether, the results show that trend inflation is more strongly related to asymmetry in responses to large shocks, than to small ones. Furthermore, I find evidence for the reverse effects of large positive shocks on production, as predicted by the model. Most importantly, these reversals are affected by trend inflation, i.e., the size of a positive shock leading to zero production response is substantially smaller in industries with high trend inflation than in those with low trend inflation. Even though these results can not be interpreted in a causal sense, they show that many of the model predictions

Figure 1.3.4: Non-Linear Cumulative Impulse Responses of Industrial Production



Impulse responses of industrial production at 6-, 12-, and 24-months horizons. Shock values are on the x-axis, measured in standard deviations. Solid green lines: industries with trend inflation above the median, dashed yellow lines: below the median, dotted black lines: entire sample. The shaded areas correspond to 68% confidence intervals, based on Newey-West standard errors, computed by the delta method.

are in line with the data.

Robustness

I show that findings discussed above are robust to several important deviations from the baseline strategy considered so far. I briefly outline the alternatives and provide the results in Appendix A.3.4.

Alternative shock measures. In the baseline specification I use a measure of monetary policy shocks, computed by Jarociński and Karadi (2020) using high frequency identification and separating monetary policy shocks from central bank information shocks using sign restrictions. I show that the results are generally robust to alternative shocks measures, commonly used in the literature. Firstly, I consider Jarociński and Karadi (2020) shock series based on a simpler separating procedure, the so called 'poor man's sign restrictions', as well as the original shock series, computed by Gertler and Karadi (2015). Secondly, I employ other widely used high-frequency identified shocks, estimated by Barakchian and Crowe (2013) and Nakamura and Steinsson (2018). Figure A.3.6 shows that the main results of the paper are in general robust to these alternative shock measures.

Measurement error in trend inflation. I estimate trend inflation at the sector level by an average PPI growth rate, which can be contaminated by a measurement error. However, I only use trend inflation to classify sectors into below and above median groups. Thus, the only way measurement error might affect the results is by distorting the ordering of sectors by trend inflation and leading to misclassification. To address this issue, I omit the middle 40% of sectors and compare the top 30% with the bottom 30%. Because sectors with trend inflation that is close to the median are much more likely to be misclassified, excluding them alleviates the problems associated with measurement error. Figures A.3.7 and A.3.8 show that results are robust to such a split.

Great Recession and ZLB. The baseline sample spans the period between February 1990 and January 2013, which includes the apex of the Great Recession and the subsequent period of low interest rates. As a robustness check, I consider a shorter sample period ending in June 2008. Figures A.3.9 and A.3.10 show that excluding the Great Recession and the ZLB period only strengthens the main results of paper.

Trimming the data. In the baseline scenario I omit the top and bottom 2.5% of sectors in terms of trend inflation from the original sample to control for potential outliers. This choice does not affect the results of the paper, with Figures A.3.11 - A.3.14 showing that the main findings are robust to trimming the top and bottom 15%, as well as to using the entire sample.

Polynomial degree. When testing for non-linearity of industrial production responses, I use a third order polynomial because it is the minimal degree that allows for non-monotonicity of impulse responses with respect to both positive and negative shocks. As a robustness check, I provide the results for the 4th, 5th and 6th order polynomials in Figure A.3.15, which shows that the effect of trend inflation on responses to large shocks does not depend on the degree of a shock polynomial.

Other. Finally, I set the number of lags to 3 and 12 (baseline specification has 6 lags) and reduce the set of controls, only keeping a time trend and lags of the dependent variable, monetary shock, and effective federal funds rate. In addition, I provide the unsmoothed impulse responses. Figure A.3.16 shows that the main findings remain unchanged.

1.3.2 Drift and Micro Level Asymmetry

In this section I show that trend inflation induces asymmetry in individual price adjustments, as follows from Proposition 1.[17] Working with price ad-

[17] Alvarez et al. (2019) work with Argentinian micro-level price data and study the effect of inflation on price behavior. The main distinction of my work is that I focus on

justments is a challenge because only continuous tracking of the price of an item can ensure that one observes adjustments as opposed to growth rates, which can consist of multiple adjustments. In addition, growth rates are functions of both adjustment size and frequency, so that any observed asymmetry in growth rates can be driven by the asymmetry in adjustment frequencies.

To address these issues, I use scraped daily data from the Billion Prices Project by Cavallo (2018). Under the assumption that prices do not change more than once a day, daily data provides the desired price adjustments. This assumption is much milder compared to those required for monthly or even weekly data. I focus on U.S. supermarket data (store 1), as it provides the longest time series, and consider items with at least 2 years of observations and at least 10 price adjustments. In addition, I exclude items that have adjustments larger than 50% to control for the outliers. The sample period is between May 2008 and July 2010, and the total number of items used in the analysis is 1924, with 28808 observed price adjustments. Figure A.3.4 in Appendix A.3.2 shows the distribution of price adjustments in the sample.

I compute asymmetry for each item i as the ratio between average sizes of positive and negative adjustments. Drift μ_i is recovered as the average price growth rate over the entire period. Baley and Blanco (2020) show that it can also be computed as the ratio between average adjustment and average time between adjustments, so I use their approach as a robustness check. The two approaches converge as the sample size increases, but can produce different estimates in finite samples. The baseline regression has the following form:

$$\log \frac{\Delta^+ p_i}{\Delta^- p_i} = \alpha_c + \beta\mu_i + \gamma' \mathbf{x}_i + \varepsilon_i$$

where $\Delta^+ p_i$ is the average positive price adjustment of item i, $\Delta^- p_i$ – average negative price adjustment, μ_i is the drift, $\mathbf{x}_i$ – a vector of controls and α_c is a category fixed effect. Items in the data are grouped into narrowly-defined categories, corresponding to the URLs where the items are found on the website. These categories are narrower than the COICOP groups and there are seven items in each category on average.[18] Including category fixed

the cross-sectional variation in item-level trend inflation, whereas they use time variation in aggregate levels of inflation. Alvarez et al. (2019) find that asymmetry in adjustments is insensitive to inflation at low inflation rates, but is positively related at high levels of inflation. I work with U.S. data and find evidence for the positive relationship even at low levels of trend inflation. A potential reason for the differences in our findings is that I consider trend inflation, i.e., the long-term growth rate of the price level, whereas Alvarez et al. (2019) focus on period-specific actual inflation, i.e., log-difference in price levels between two consecutive periods.

[18] I exclude categories with less than 3 items to allow for enough within-category variation.

effects controls for many unobservables such as category-specific demand, adjustment costs, or other characteristics that may simultaneously affect both the drift and the asymmetry. The set of controls includes the frequency and standard deviation of adjustments, as well as the variance of idiosyncratic shocks σ_i^2, computed following Baley and Blanco (2020). All variables are normalized to monthly frequency, and summary statistics are provided in Table A.1 in Appendix A.3.2.

The first three columns of Table 1.1 show the results from an OLS regression with standard errors clustered at the category level. Columns (1) and (2) employ the baseline estimates of μ_i as an average price growth rate, and an alternative measure for μ_i (as in Baley and Blanco (2020)) is used in column (3).

Table 1.1: Micro-level Asymmetry

	Dependent variable:				
	Asymmetry $\left(\log \frac{\Delta^+ p_i}{\Delta^- p_i}\right)$				
	OLS			IV	
	(1)	(2)	(3)	(4)	(5)
Drift μ	4.969***	4.966***		11.407**	
	(1.830)	(1.873)		(5.448)	
Drift μ (alt.)			4.552***		36.707
			(1.745)		(25.146)
σ^2		−0.899	−0.890		
		(3.173)	(3.175)		
Frequency		0.114	0.113		
		(0.151)	(0.151)		
Std. Dev.		0.028	0.024		
		(0.730)	(0.730)		
Observations	1,924	1,924	1,924	1,376	1,376
R^2	0.458	0.460	0.460	0.483	0.413

Note: *p<0.1; **p<0.05; ***p<0.01. All specifications include category FE. Standard errors are clustered at category level.

Table 1.1 suggests that higher trend inflation is positively related to asymmetry in individual adjustments, independent of the way drift μ is computed. The inclusion of controls does not alter this result. The coefficient in the first column is interpreted in the following way: a one percentage point increase in monthly trend inflation is associated with a 5% increase in the size of positive adjustments relative to the size of negative adjustments. Note that a 1 p.p. increase in trend inflation is not a very large change at the item level: standard deviation of the drift μ_i is 0.8 p.p. in the cross-sectional distribution, so that the drift effect is sizable.

As noted previously, the positive relationship between the average growth rate and the asymmetry in adjustments may not be too surprising, but it is not immediate either. A higher trend may be purely driven by more frequent positive adjustments and less frequent negative adjustments, however, this option is not supported by the data.

One potential drawback of the baseline OLS specification is the fact that drifts and asymmetries are computed using the same item-level time series. This may lead to spurious results in a short sample because a large positive adjustment simultaneously increases the estimates of drift and asymmetry. To resolve this issue, I split the sample into two equal parts for each item. I use the drift in the first subsample as an instrument for the drift in the second subsample. I then compute asymmetry in the second subsample and regress it onto the instrumented drift. Thus, the drifts and asymmetries are effectively estimated on different samples, which helps addressing this issue. The results are presented in columns (4) and (5) of Table 1.1. The coefficient in front of the drift increases, and so do the standard errors.[19] The baseline estimate of the drift remains significant and the alternative specification becomes marginally insignificant with p-value = 0.14. Overall, I conclude that the results are robust and provide supporting evidence for the model predictions regarding the link between trend inflation and asymmetry at the level of individual price adjustment.

1.4 Monetary Policy in General Equilibrium

The analytic results of this paper provide new insights into the efficacy of monetary policy and its dependence on trend inflation. However, these results are obtained in a rather restrictive environment. First, I assume that firms follow the steady state optimal policy along the transition path. Second, I use a second-order approximation of the profit function, which ensures symmetry and substantially contributes to analytic tractability. Third, I consider monetary interventions in isolation, whereas this policy instrument is often used as a counteractive measure to mitigate the effects of other disturbances. Therefore, monetary policy is often implemented outside of an

[19] The standard errors increase due to the instrumenting procedure and a smaller sample as I additionally restrict attention to items with at least 5 adjustments in the second subsample.

economy's steady state, in contrast to the assumption imposed in the analytic framework.

I now address all these issues and embed the analytic framework into a standard general equilibrium model, calibrated to the U.S. data. I consider a transitory adverse markup shock, which leads to an increase in prices and a drop in consumption. Firms now correctly anticipate the economy dynamics and follow the appropriate optimal policy. I then compare the ability of a monetary authority to stabilize the economy under the baseline 2% inflation target and a counterfactual 4% inflation target. A markup shock is well-suited for this exercise, as it only increases a wedge in the economy stemming from price dispersion, without affecting the efficient allocation. This provides a rationale for the imposed stabilization objective of the monetary authority. Because the markup shock depresses consumption and increases prices, it introduces a trade-off for the monetary authority, as it can not stabilize consumption and prices simultaneously.

I find that increasing the inflation target imposes two negative effects on the ability of a policymaker to stabilize the economy after such a shock. First, higher trend inflation amplifies the initial effect of the markup shock, leading to larger price and consumption deviations. Second, it worsens the trade-off between price and consumption stabilization. Both effects are sizable and arise due to the impact of trend inflation on the asymmetry of price and output responses. The results relate to the ongoing discussion on increasing the inflation target, highlighting adverse implications for stabilization policy away from the zero lower bound, in particular for the type of shocks that exhibit the 'cost-push' property of moving prices and output in opposite directions.

1.4.1 General Equilibrium Setup

Households

I embed the analytic model from Section 1 into a general equilibrium setting, similar to those in Nakamura and Steinsson (2010) and Karadi and Reiff (2019). Representative households maximize the present discounted value of their utility, given by

$$\int_0^\infty e^{-\rho t} \left(\log C_t - \alpha L_t\right) dt$$

where C_t denotes consumption of a composite good, L_t is the household's labor supply, ρ is the discount rate, and α is the disutility of labor. The household's budget constraint is as follows:

$$P_t C_t + \dot{B}_t = R_t B_t + W_t L_t + \Pi_t$$

where P_t is the aggregate price level, B_t are the holdings of a bond with nominal gross return R_t, W_t is the wage and Π_t are the firms' profits. Consumption C_t is composed of a continuum of differentiated goods and is given by

$$C_t = \left[\int \left(A_t(i) C_t(i) \right)^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}}$$

where $C_t(i)$ is consumption of a good produced by firm i, $A_t(i)$ is its quality, and θ is the elasticity of substitution. The aggregate price level is $P_t = \left[\int \left(P_t(i)/A_t(i) \right)^{1-\theta} di \right]^{\frac{1}{1-\theta}}$ and cost minimization yields the following demand for good i:

$$C_t(i) = A_t(i)^{\theta-1} \left[\frac{P_t(i)}{P_t} \right]^{-\theta} C_t$$

First-order conditions imply that wage W_t is proportional to nominal aggregate consumption $P_t C_t$, and nominal interest rate is determined by the growth rate of nominal consumption:

$$W_t = \alpha P_t C_t$$
$$R_t = \rho + \frac{(\dot{P_t C_t})}{P_t C_t}$$

Firms

There is a continuum of firms producing differentiated goods, indexed by $i \in [0, 1]$. Firms demand labor $L_t(i)$ and set prices $P_t(i)$. Production technology is given by $Y_t(i) = L_t(i)/A_t(i)$, so that higher quality goods are more costly to produce. Firms' profits are given by $\Pi_t(i) = P_t(i) Y_t(i) - W_t L_t(i)$. To adjust its price at time t, a firm must hire additional labor and the total cost of adjustment is given by $\kappa P_t(i) Y_t(i)$. In addition, firms receive an opportunity to adjust for free at rate λ. Such a setup is typically referred to in the literature as a 'CalvoPlus' model because it nests both the standard menu cost model and the Calvo (1983) setting. Each firm maximizes the expected discounted stream of profits:

$$\mathbb{E} \left[\int_0^\infty Q_t \Pi_t(i) dt - \kappa \sum_{i=1}^{\infty} Q_{\tau_i} P_{\tau_i}(i) Y_{\tau_i}(i) \right]$$

where $Q_t = \frac{\alpha e^{-\rho t}}{W_t}$ is the discount factor implied by the household's problem and τ_i are the adjustment times when a firm pays adjustment costs. The goods quality $A_t(i)$ evolves as a geometric Brownian motion with no drift: $d \log A_t(i) = \sigma dW_t(i)$. Using the household's first-order conditions and the fact that firms face consumers' demand function ($Y_t(i) = C_t(i)$), one can rewrite the firm's profit and cost functions as:

$$\Pi_t(i) = \alpha^{-\theta} W_t \left(\frac{\theta C_t}{\theta - 1}\right)^{1-\theta} \overbrace{e^{-\theta z_t(i)} \left(e^{z_t(i)} - \frac{\theta - 1}{\theta}\right)}^{\pi(z_t(i))}$$

$$\kappa P_t(i) Y_t(i) = \kappa \alpha^{-\theta} W_t \left(\frac{\theta C_t}{\theta - 1}\right)^{1-\theta} \underbrace{e^{(1-\theta) z_t(i)}}_{c(z_t(i))}$$

where $z_t(i)$ is the price gap, given by $z_t(i) = \log P_t(i) - \log P_t^*(i)$, and $P_t^*(i)$ is the frictionless optimal price, given by $P_t^*(i) = \frac{\theta}{\theta - 1} W_t A_t(i)$. Note that W_t cancels out in the firm's objective function, so that in a stationary equilibrium the firm's problem does not depend on any aggregate state, as constant aggregate consumption may be taken out of the problem.

Monetary Authority and Stationary Equilibrium

Following Nakamura and Steinsson (2010) and Midrigan (2011), I assume that the monetary authority is in full control of the nominal output $M_t = P_t C_t$, which in the steady state grows at a constant rate μ: $d \log M_t = \mu dt$. This assumption is common in the literature and can be rationalized by a binding cash-in-advance constraint. Given that μ is set exogenously in this model, I will refer to it both as 'trend inflation' and 'inflation target'.

Household's first-order conditions imply that the equilibrium nominal interest rate is constant and equal to $\bar{R} = \rho + \mu$, and the wage follows the law of motion of nominal output: $d \log W_t = \mu dt$. This creates a drift in the firm's optimal price $P_t^*(i)$, and thus in the price gaps $z_t(i)$. In the absence of action, price gaps evolve as $dz_t(i) = -\mu dt + \sigma dW_t(i)$. The firm's problem becomes almost identical to the one considered in the analytic section, with a few exceptions: (1) the profit function is no longer symmetric, (2) adjustment costs depend on the price gap at the time of adjustment, and (3) firms receive costless adjustment opportunities at rate λ. The solution to the firm's problem is characterized by a triplet $\{\underline{z}, \hat{z}, \overline{z}\}$ where $\underline{z}$ and $\overline{z}$ are the lower and upper boundaries of inaction region, and $\hat{z}$ is the return point. The value function satisfies the following Hamilton–Jacobi–Bellman equation in the inaction region:

$$(\rho + \lambda)v(z) = \pi(z) + \lambda v(\hat{z}) - \mu v'(z) + \frac{1}{2}\sigma^2 v''(z)$$

where $\pi(z) = e^{-\theta z}\left(e^z - \frac{\theta-1}{\theta}\right)$ and $\hat{z}$ is the optimal return point. The boundary conditions are $v(\underline{z}) = v(\hat{z}) - c(\underline{z})$ and $v(\overline{z}) = v(\hat{z}) - c(\overline{z})$, where $c(z) = e^{(1-\theta)z}$. Optimality and smooth pasting require $v'(\hat{z}) = 0$, $v'(\underline{z}) = (\theta - 1)c(\underline{z})$ and $v'(\overline{z}) = (\theta - 1)c(\overline{z})$. The density of the stationary price gap distribution $f(z)$ is determined by a Kolmogorov forward equation:

$$\lambda f(z) = \mu f'(z) + \frac{1}{2}\sigma^2 f''(z)$$

Aggregate consumption, price level and employment can be computed using the stationary price gap distribution as follows:

$$C_t = \bar{C} = \frac{\theta - 1}{\alpha\theta}\left[\int_{\underline{z}}^{\overline{z}} e^{(1-\theta)z} f(z)dz\right]^{\frac{1}{\theta-1}}$$

$$P_t = \frac{\alpha\theta}{\theta - 1} M_t \left[\int_{\underline{z}}^{\overline{z}} e^{(1-\theta)z} f(z)dz\right]^{\frac{1}{1-\theta}}$$

$$L_t = \bar{L} = \bar{C}^{1-\theta}\left(\frac{\alpha\theta}{\theta-1}\right)^{-\theta} \int_{\underline{z}}^{\overline{z}} e^{-\theta z} f(z)dz \; + \; \Gamma$$

$$\Gamma = \kappa\alpha^{-\theta}\left(\frac{\theta\bar{C}}{\theta-1}\right)^{1-\theta}\left[\gamma^+ e^{(1-\theta)\underline{z}} + \gamma^- e^{(1-\theta)\overline{z}}\right]$$

where Γ is the total labor hired for price adjustment per unit of time and γ^+ and γ^- are the masses of firms adjusting at a cost upward or downward, respectively, per unit of time. Finally, bond holdings B_t are in zero net supply, so that in equilibrium $B_t = 0$.

1.4.2 Calibration

I set the discount rate ρ to 0.04 in annual terms and trend inflation μ to 0.02, roughly matching the average annual inflation in the U.S. over the last two decades.[20] The elasticity of substitution θ is set to 5, which is an intermediate value among those considered in the literature.[21] The remaining parameters,

[20] When calibrating a continuous time model, the period length is innocuous, as it only scales certain parameters up or down.

[21] Midrigan (2011) sets θ to 3, Nakamura and Steinsson (2010): $\theta = 4$, Karadi and Reiff (2019): $\theta = 5$, Golosov and Lucas (2007) use $\theta = 7$.

namely the disutility of labor α, the variance of idiosyncratic shocks σ^2, the adjustment cost κ, and the rate at which firms receive free adjustment opportunities λ, are calibrated internally. I target equilibrium employment of 1/3 and three moments of the distribution of price adjustments: frequency, average size, and kurtosis. All three moments are informative of aggregate responses to shocks and are a typical choice for calibration targets. Alvarez et al. (2016) show analytically that in a wide class of menu cost models the ratio of kurtosis to frequency is a sufficient statistic for the cumulative effect of a marginal monetary shock on output. In the first section of this paper I show that the effect of trend inflation on aggregate price and output responses depends on the average size of adjustment.

I target values of frequency, average size and kurtosis, reported in the literature. I set the frequency of price changes to 10% per month, the average size of adjustment to 10%, and the kurtosis of the distribution of price adjustments to 4. The first two values are standard, as many studies report very similar estimates using different data sets.[22] The estimates of kurtosis are much more dispersed: Alvarez et al. (2020) report values close to 2, Midrigan (2011): 3.15, Alvarez et al. (2016): 4, Vavra (2014): 6.4. I use an intermediate value of 4, obtained by Alvarez et al. (2016) from the weekly scanner data of the Dominick's dataset, accounting for heterogeneity and measurement error. The model matches the targeted statistics exactly, and Table 1.2 summarizes model parameters and their values in annual terms. In Appendix A.4.3 I use an alternative calibration, targeting the kurtosis of the price adjustment distribution of 3. This affects the overall non-neutrality of monetary policy, but the main findings remain qualitatively unchanged.

1.4.3 Markup Shock

I now consider an unexpected shock that increases steady state optimal markup $\left(\frac{\theta}{\theta-1}\right)$ by 3% and then gradually reverts to zero in AR(1) fashion. Formally, the dynamics of the shock ε_t are governed by an Ornstein-Uhlenbeck process, so that $\varepsilon_t = 0.03 \cdot e^{-\eta t}$, where η determines the speed of convergence and is set to generate a half-life of two months. The shock sets the economy on a deterministic transition path, increasing the aggregate price and depressing consumption. I defer the description of the non-

[22] Frequency: Nakamura and Steinsson (2008): 10.8%, Nakamura and Steinsson (2010): 8.7%, Vavra (2014): 10.9%. Average size of adjustment: Nakamura and Steinsson (2008): 8.5%, Kehoe and Midrigan (2015): 11%, Vavra (2014): 7.7%. For the average size of adjustment, the mean and median estimates are usually similar, whereas the mean frequency is typically higher than the median. I use the median frequency estimates, as this is the preferred choice for single-sector models (see Nakamura and Steinsson (2010)).

Table 1.2: Calibrated Model Parameters

Parameter	Value
Discount rate ρ	0.04
Trend inflation μ	0.02
Elasticity of substitution θ	5
Disutility of labor α	2.2
Variance of idiosyn. shocks σ	0.148
Adjustment cost κ	0.11
Rate of free adjustmens λ	1.126

Values are denominated in annual terms.

stationary equilibrium conditions to Appendix A.4.1 and plot the dynamics of consumption and prices on Figure 1.4.1.

The price level response is plotted in terms of percent deviations from the trend, whereas consumption and markup responses are in terms of percent deviations from the steady state. The markup shock raises the firms' optimal prices, leading to an increase in the actual price level. Because the nominal output stays constant and prices increase, consumption falls. Integrating the area under the lines, one obtains cumulative impulse responses, which are given by $\int_0^\infty (p_t - \bar{p}_t)dt = 0.44\%$ for the price level and $\int_0^\infty (c_t - \bar{c})dt = -0.44\%$ for consumption, where $\bar{p}_t$ is the trend of the aggregate log-price and $\bar{c}$ is the steady state log-consumption.[23]

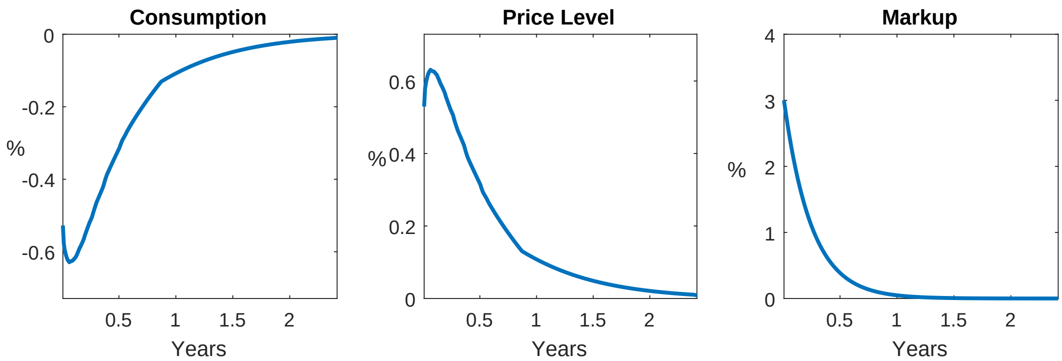
The shock is purely inefficient in the sense that it increases the wedge between the actual and efficient level of output, without affecting the efficient allocation.[24] Thus, it would be desirable to 'undo' its consequences by means of policy. To capture this in a simple way, I assume that the policymaker dislikes negative deviations of consumption from its efficient level and values price stability (dislikes any deviations from the trend).[25] I also assume that monetary interventions follow the same dynamics as the markup shock, and the only choice of the policymaker is the level of monetary inter-

[23]A 1% negative cumulative response of consumption is equivalent to a scenario when consumption is held at 1% below its steady state for one year.

[24]Efficient output is achieved under zero price dispersion and is given by $C_t^* = L_t$. Due to price dispersion, $C_t = \left[\int (P_t(i)/(A_t(i)P_t))^{-\theta}\, di\right]^{-1} C_t^*$ (see Yun (1996)). An increase in optimal markup lowers θ and increases the inefficiency stemming from price dispersion.

[25]Such an objective is different from an optimal policy that considers welfare, which depends on the level of consumption, the degree of price dispersion and the volume of adjustment costs paid by the firms. A meaningful study of optimal policy with respect to trend inflation would require additional model components, e.g. heterogeneity in individual price trends as in Adam and Weber (2020).

Figure 1.4.1: Markup Shock



Model-generated impulse responses of consumption, price level and markup to a 3% markup shock. Consumption and markup responses are in terms of percent deviations from the steady state, whereas price level responses are in terms of percent deviations from the trend.

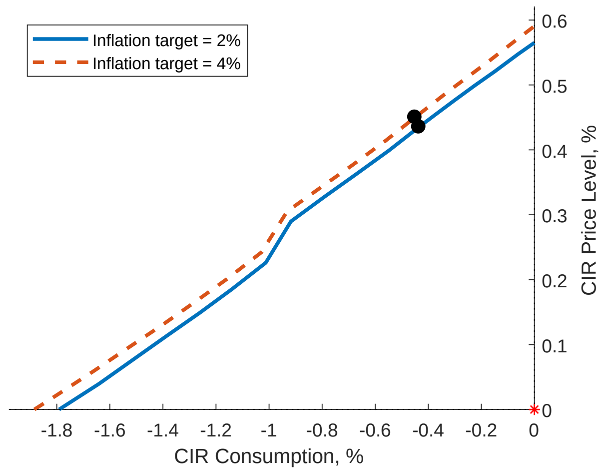
vention. Formally, monetary intervention δ_t is proportional to the markup shock: $\delta_t = \delta\varepsilon_t$, where $\delta \in \mathbb{R}$ and is chosen by the monetary authority. A stimulus ($\delta > 0$) mitigates the negative response of consumption, but raises prices even further (see Figure A.4.1 in Appendix A.4.2). A contraction ($\delta < 0$) creates an opposite effect, stabilizing prices and amplifying the drop in consumption. The policymaker thus faces a trade-off, as it is impossible to stabilize consumption and prices simultaneously.

I do not assign any weights to these objectives, but rather consider the whole possibility frontier of the policymaker, given the initial markup shock and the restrictions on policy outlined above. By varying the sign and size of the monetary intervention δ, the policymaker achieves different combinations of cumulative consumption and price responses. The resulting frontier depends, among other parameters, on trend inflation μ. I now compare these frontiers for the baseline level of trend inflation of 2% per year and a counterfactual value of 4%. Figure 1.4.2 shows the results.

On the x-axis I plot cumulative consumption responses, on the y-axis – cumulative responses of the price level.[26] The curves show feasible outcomes for the baseline economy with trend inflation of 2% (solid blue line) and a counterfactual economy with a 4% trend inflation (dashed red line), given the

[26] I consider cumulative deviations of the price level rather than inflation, as in this case it is impossible to completely neutralize the effect of the shock on inflation due to the imposed restriction on monetary policy. However, the findings of the paper remain unchanged if I substitute the price level CIR with the CIR of inflation, as shown in Appendix A.4.4.

Figure 1.4.2: Frontiers, Small Shock



Feasible combinations of cumulative responses of consumption (x-axis) and price level (y-axis) after a 3% markup shock. The solid blue line corresponds to the baseline economy with a 2% trend inflation, and the dashed red line represents a counterfactual economy with a 4% trend inflation. Consumption responses are in terms of percent deviations from the steady state, whereas price level responses are in terms of percent deviations from the trend. Black dots show the outcomes if the monetary authority does not intervene.

initial 3% markup shock. The red asterisk corresponds to a (0, 0) scenario, where the effect of the markup shock is completely neutralized. Black dots on the curves correspond to scenarios when the monetary authority does not intervene ($\delta = 0$). Stimulative policy ($\delta > 0$) moves an economy along its frontier to the right, contractionary measures ($\delta < 0$) move it to the left.

First, note that in the economy with a 4% trend inflation the black dot is further away from the (0, 0) point, which means that the negative effects of the markup shock are on their own stronger if trend inflation is higher. Under a 4% inflation target, the markup shock leads to a 3.4% stronger increase in prices and a 3.4% stronger drop in consumption, compared to the baseline economy with an inflation target of 2%. When trend inflation is higher, prices exhibit less upward rigidity and have a stronger response to the markup shock, which also results in a larger consumption drop if the monetary authority keeps the nominal output constant. Importantly, this result is not driven by changes in the overall frequency of price adjustments because it remains virtually constant as I vary the level of trend inflation. Instead, the result is due to changes in the relative frequencies and sizes of positive and negative price adjustments.

Second, higher trend inflation worsens the trade-off between consumption and price stabilization. This latter effect is less apparent on the graph, but can be seen when calculating the curvature of the frontiers. I measure the curvature as a ratio between the slopes of stimulative and contractionary interventions. The slope of stimulative policy α_S reflects the rate at which the policymaker gains consumption and loses price stability, when conducting stimulative policy ($\delta > 0$). Graphically, it is the slope of a straight line, passing through an economy's initial point (black dot) and the intersection of the frontier with the y-axis. The slope of contractionary policy α_C reflects the rate at which the policymaker gains price stability and loses consumption, when conducting contractionary policy ($\delta < 0$). Graphically, it is the slope of a straight line, passing through an economy's initial point (black dot) and the intersection of the frontier with the x-axis. The curvature is then measured as a ratio between the slopes: α_S/α_C.[27] A higher curvature indicates that the stimulative slope becomes steeper, whereas the contractionary slope flattens out. Therefore, the monetary authority must sacrifice more consumption when stabilizing prices and tolerate larger price deviations when restoring consumption. Thus, the higher the curvature, the worse the stabilization trade-off is.

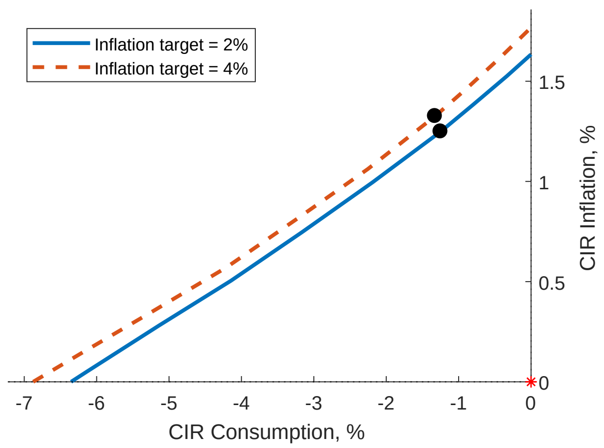
For the baseline economy with a 2% inflation target the curvature is equal to 0.93, whereas under a 4% inflation target it increases by by 7.5% to 1.0. Under higher trend inflation, the policymaker must sacrifice more consumption when stabilizing prices, and must tolerate larger price responses when stimulating consumption. This is again caused by the effect of trend inflation on the asymmetry of price and consumption responses. As the inflation target rises, prices become more sensitive to stimulative shocks and it becomes harder for the monetary authority to stimulate consumption. Simultaneously, prices become less sensitive to contractionary shocks, which impedes the ability of policymakers to stabilize prices. Higher trend inflation increases price flexibility exactly when it is desirable to have rigid prices, and makes them stickier exactly when flexibility is needed.

Both of the effects of a higher inflation target are amplified if the initial markup shock is larger. Figure 1.4.3 plots the same frontiers for a 10% markup shock. The economy with a 4% trend inflation now has a 6.1% stronger response to the initial markup shock and an 11% higher curvature, compared to the economy with a 2% trend inflation.

Overall, the results show that trend inflation affects the ability of a policymaker to stabilize the economy after an adverse markup shock. Higher

[27] This measure is not ideal, as it assigns a unique value to the entire frontier, whereas the degree of curvature may vary along the frontier. However, it summarizes the overall trade-off, considering two extreme points of achieving zero consumption or zero price CIRs.

Figure 1.4.3: Frontiers, Large Shock



Feasible combinations of cumulative responses of consumption (x-axis) and price level (y-axis) after a 10% markup shock. The solid blue line corresponds to the baseline economy with a 2% trend inflation, and the dashed red line represents a counterfactual economy with a 4% trend inflation. Consumption responses are in terms of percent deviations from the steady state, whereas price level responses are in terms of percent deviations from the trend. Black dots show the outcomes if the monetary authority does not intervene.

trend inflation decreases upward price stickiness and leads to stronger price and consumption responses to the initial markup shock. In addition, higher trend inflation amplifies the asymmetry of price and consumption responses to positive and negative monetary shocks, which worsens the policymaker's trade-off when stabilizing the economy. I finally note that these results are of greater importance for large shocks, as the effects of trend inflation become more pronounced.

In both scenarios I considered a shock that temporarily increases monopolistic power of firms. A shock that decreases firms' monopolistic power and drives optimal markups down would have two distinct effects. First, because the steady state markup is positive and price dispersion is non-zero, a fall in markups would decrease the inefficiency in the economy and bring consumption closer to its efficient level. Such a shock would increase consumption and decrease prices, so that any subsequent expansionary monetary policy would lead to price stabilization and further consumption growth, thus inducing no trade-off. Second, because higher trend inflation increases downward price rigidity, the initial response to the shock would be larger in the baseline economy with a 2% inflation target than in the counterfactual with a 4% inflation target. In addition, it will be easier for the monetary authority to stabilize prices and consumption under higher trend inflation, again due to lower upward price rigidity and higher downward rigidity. Therefore, all results are 'mirrored' if the markup shock is of the opposite sign, and higher trend inflation would be beneficial from a pure stabilization perspective.[28] It follows that the overall potency of monetary stabilization policy would depend on which types of shocks prevail in the economy.

1.5 Summary

In this paper I show that trend inflation matters for economy's responses to aggregate shocks and monetary policy interventions. I derive a set of new analytic results for the effect of trend inflation on aggregate dynamics in a standard menu cost model. The main contribution is that I consider monetary shocks of any size in an environment with non-zero drift. This approach reveals several new properties of aggregate dynamics, especially for large shocks.

The key characteristic of trend inflation is that it affects aggregate responses to positive and negative shocks asymmetrically. In the presence of adjustment costs, prices are more sensitive to shocks that push them in the same direction as the trend, and are less sensitive to shocks the push them in the opposite direction. Under positive trend inflation, larger price flexibility in responses to positive monetary shocks leads to weaker output increases, whereas smaller price flexibility in responses to negative shocks leads to stronger output declines. These effects are especially pronounced for large shocks that force all firms to update prices. While positive large shocks are neutral in the driftless case, they cause output contractions in economies with positive trend inflation.

The empirical analysis shows that the new analytic predictions of the model are in line with the data. I find that sectors with a higher PPI growth rate exhibit stronger price responses to positive monetary shocks and weaker responses to negative shocks, compared to sectors with a lower growth rate of PPI. I also find that aggregate output expansions after positive monetary shocks are almost entirely driven by sectors with a low PPI growth rate, whereas output contractions are distributed more equally. In addition, production responses are generally non-linear and large positive shocks may lead to a decline in output. This holds for sectors with both low and high levels of trend inflation, however the size of a positive shock that causes an output contraction is smaller for sectors with a larger level of trend inflation.

[28] See also the discussion in Blanco (2020) on the effects of higher trend inflation on the likelihood of hitting the zero lower bound.

My results have important implications for monetary stabilization policy and contribute to the ongoing discussion on the necessity to raise the inflation target. Using a general equilibrium model calibrated to the U.S. data, I find that higher trend inflation has a sizable effect on the ability of a policymaker to stabilize the economy after an adverse markup shock. Raising the inflation target from 2% to 4% amplifies the initial response to the markup shock and worsens the stabilization trade off. A policymaker has to sacrifice more consumption when stabilizing prices and has to tolerate larger price deviations when stimulating consumption. Thus, a higher inflation target impedes the ability of a monetary authority to counteract adverse shocks that move output and prices in opposite directions.

Chapter 2

Understanding Leverage Determinants

2.1 Introduction

Secured loans are a common type of borrowing, in which the borrower pledges an asset as collateral to ensure the lender against potential default. This allows the borrower to negotiate a lower interest rate and obtain a larger loan. In many cases, agents borrow to purchase an asset and use that asset to secure the loan. One example are mortgages: households borrow to buy a house and use the house as collateral. A key statistic that affects the ability of households to purchase houses is leverage. Leverage is the ratio between the value of an asset (the price of a house) and agent's equity (household's down payment). If a household buys a house borrowing 80% of its price, then the down payment amounts to 20% of the price, and the leverage is equal to 5. The margin is the reciprocal of the leverage, and in this example it is equal to 20%.

Leverage on secured loans experienced violent fluctuations around the Great Recession. The procyclicality of leverage, documented by Adrian and Shin (2010), was exceptionally pronounced during this period. According to Geanakoplos (2010), the average leverage on so-called toxic mortgage backed securities went down from 16 in 2006 to 1.2 in 2009. Within the same period, the down payment on mortgages went up from as low as 3% to 30%. Fluctuations originating in the financial sector spilled over to the real economy and resulted in a severe recession. The aim of this paper is to provide new theoretical insights into forces and mechanisms responsible for such drastic movements.

To understand the determinants of leverage, I setup a general equilibrium model with an endogenous leverage constraint. The model features a risky asset with a stochastic future payoff and a safe asset. Besides investing in either of the two assets, agents can borrow from each other. Borrowing, however, requires pledging one unit of the risky asset as collateral, and the payoff of the risky asset can not be influenced by the borrower or the lender. Examples of such borrowing contracts are REPO loans and mortgages. The agents disagree about the distribution of the payoff of the risky asset, meaning that some are more optimistic than the others. In equilibrium, the optimists would like to buy the risky asset and to borrow from the more pessimistic agents to increase their asset purchases. However, the value of the risky asset that is used as collateral differs across borrowers and lenders. The lenders are more pessimistic and require larger amounts of collateral for a given loan. This creates an endogenous leverage constraint for the borrowers.

The model is based on the framework developed by Geanakoplos (1997) and extended by Simsek (2013). The novelty of my approach is that I simultaneously consider a continuum of agent types and a continuum of states for the payoff of the risky asset. A continuum of agent types (as opposed to two types in Simsek (2013)) ensures that identities of borrowers and lenders are determined endogenously in equilibrium. A continuum of states for the risky asset payoff (as opposed to two or three in Geanakoplos (2003)) ensures that agents with different beliefs choose different borrowing contracts in equilibrium. Such a setup is not a mere technical complication, but it provides new theoretical insights, inaccessible otherwise. The model features a continuum of different borrowing contracts traded in equilibrium, all written against the same asset used as collateral. There is one-to-one matching between borrowers and lenders and each pair has a unique contract in terms of riskiness, leverage and promised interest. The most optimistic agents borrow with the highest leverage and promise the largest interest rates. An important new analytic result is the absence of riskless borrowing contracts in equilibrium: all contracts default if the realized asset payoff is sufficiently low.

The model also highlights that the price of the risky asset and the total leverage in the economy are driven by different forces. The asset price is determined by the ratio between the mass of market participants (those buying the risky asset or lending) and the rest of the population. The more agents participate in the market, the larger fraction of total endowment is invested into the asset, and the higher is the asset price. The aggregate leverage, in turn, depends on the ratio between the mass of lenders and the mass of agents buying the asset (borrowers). If the mass of lenders increases relative to the mass of borrowers, the leverage goes up. Therefore, the aggregate leverage and the asset price are decoupled: there can be more agents participating in the market and raising the asset price, but as long as these agents split into borrowers and lenders in the same proportion, the leverage stays constant. Similarly, agents may be switching from borrowing to lending and thus increasing the leverage, but as long as the total mass of market participants stays constant, so does the asset price.

To see which fundamentals have the largest effect on leverage, I solve the model numerically and compute comparative statics with respect to two parameters: the optimism and the uncertainty. The optimism is the average (across agents) expected payoff of the risky asset. The uncertainty is the variance of this payoff, and it is constant across agents. I find that aggregate leverage is much more sensitive to changes in the uncertainty, than to changes in the optimism. Higher uncertainty scares lenders away, and they either flee the market or switch to buying the asset on margin instead of lending. As a result, the mass of lenders goes down, whereas the mass of borrowers increases, and aggregate leverage falls. The reason is that returns on lending are sensitive to changes in the lower tail of the asset payoff distribution, whereas returns on buying the asset on margin primarily depend on the upper tail of the distribution. As uncertainty rises, the upper tail 'improves' in the sense that larger asset payoffs become more likely, whereas the lower tail 'worsens' in the sense that smaller asset payoffs also become more likely. As a result, more agents are willing to buy the asset on margin and fewer agents are willing to lend, which leads to lower aggregate leverage.

Changes in the optimism, however, primarily affect the price of the asset. Higher optimism attracts more agents on the market, which increases the asset price. At the same time, I find the effect on leverage to be small and ambiguous. As the number of market participants rises, the fractions of borrowers and lenders remain relatively stable, and so does the aggregate leverage.

Relation to the Literature. This paper is most closely related to the literature that studies the role of collateral in general equilibrium models with incomplete markets (GEI). The pioneering work of Geanakoplos (1997, 2003) shows that constraints on leverage arise endogenously due to scarcity of the collateral and belief heterogeneity. Geanakoplos and Zame (2014) establish the main properties of collateral equilibrium. Fostel and Geanakoplos (2008) and Geanakoplos (2010) introduce the notion of the leverage cycle, in which an increase in uncertainty leads to a spiral of falling leverage and asset prices. Fostel and Geanakoplos (2015) show that in models with a continuum of agent types and only two states of nature, only the no-default borrowing contract is traded in equilibrium. Simsek (2013) considers an opposite setting, allowing for a continuum of states but only two agent types, and explores how the nature of belief disagreement affects equilibrium outcomes. The most recent contributions follow the approach of Simsek (2013). Yan (2017) studies a version of this model with no risk-free asset and consumption in both periods. Pei and Zhang (2020) incorporate this framework into an infinite horizon RBC model and employ heterogeneity in productivity instead of belief heterogeneity. I contribute to the literature by combining a continuum of states with a continuum of agent types. This allows me to derive new theoretical predictions, unattainable in other environments, such as the absence of no-default borrowing contracts in equilibrium, in contrast to the Binomial No-Default Theorem of Fostel and Geanakoplos (2015).

My paper also fits into broader literature that investigates the effects of financial frictions on asset prices and leverage, including Shleifer and Vishny (1992), Kiyotaki and Moore (1997), Bernanke et al. (1999), and more recently Brunnermeier and Pedersen (2009), He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014), and Coimbra and Rey (2020). The key distinctive features of the setting in my paper are the type of the financial friction and the source of the endogenous constraint on leverage. The collateral constraint in my setup does not apply to the borrowed amount, but to the number of borrowing contracts sold by an agent. This implies that agents may borrow more against the same amount of collateral and increase leverage by promising a higher interest rate. The menu of borrowing contracts is an equilibrium outcome and specifies available combinations of leverage and interest rates. As a result, the constraints on leverage arise endogenously and are driven by the difference in the valuation of the collateral between borrowers and lenders.

Structure of the paper. The next section outlines the model setup and defines the general equilibrium. Section 3 characterizes the structure of the equilibrium. Section 4 establishes an alternative equilibrium definition. Section 5 computes comparative statics. Section 6 concludes and suggests avenues for future research.

2.2 Model Setup

I consider an economy populated by a continuum of agents $i \in [0, 1]$. There are two periods, $t = 0$ and $t = 1$, and agents consume in $t = 1$ only. There are two commodities: a risk-free asset, delivering one unit of consumption good in $t = 1$, and a risky asset, delivering $y \in [\underline{c}, \overline{c}]$ units of consumption good in $t = 1$, with $\underline{c} > 0$. Following Simsek (2013), I will refer to the risk-free asset as 'cash', and to the risky asset simply as 'asset'. Both commodities are traded in period $t = 0$. The price of cash is normalized to 1, which sets the gross risk-free rate to 1. The price of the asset in $t = 0$ is denoted by p and is determined endogenously. There is one unit of each commodity, distributed equally among agents, so that agent's endowment is equal to $1 + p$.

Agents have linear utility and heterogeneous beliefs about the distribution of the asset payoff in $t = 1$, represented by cumulative distribution functions $F_i : [\underline{c}, \overline{c}] \to [0, 1]$. These distributions have continuous densities $f_i(\cdot)$, which are positive over $(\underline{c}, \overline{c})$. The main source of belief heterogeneity is optimism, represented by the expected value of the asset payoff $\mathbb{E}_i[y]$. Agents are ranked by their optimism, in the sense that agents with a higher index expect a higher asset payoff, formally: $\forall\, i, j \in [0, 1]$ s.t. $i < j$, $\mathbb{E}_i[y] < \mathbb{E}_j[y]$. The most optimistic agents belong to the top of the $[0, 1]$ interval, the most pessimistic – to the bottom.

In equilibrium, the most optimistic agents would hold the asset, and the most pessimistic ones would invest in risk-free cash. Furthermore, optimists would like to borrow in order to purchase more of the asset.[1] In other words, optimists would prefer to buy the asset 'leveraged' or 'on margin', since borrowed funds are used to purchase more of the asset. In the paper I use the two terms interchangeably.

I model borrowing in a form of anonymous market for borrowing contracts, as in Geanakoplos (2003, 2010). A borrowing contract is a promise to deliver φ units of consumption good in $t = 1$, collateralized by one unit of the asset. Contracts are traded in $t = 0$ at endogenously determined prices $q(\varphi)$. By selling a contract, an agent puts up one unit of the asset as collateral, borrows $q(\varphi)$ in $t = 0$ and promises to deliver φ in $t = 1$, implying a gross interest rate of $\varphi/q(\varphi)$. By buying a contract, an agent lends $q(\varphi)$ in $t = 0$ and is promised φ in $t = 1$.

However, a promise might not be delivered if the borrower prefers to default. This happens if fulfilling the promise is more costly than surrendering the collateral, namely if the promised amount φ is larger than the value of collateral y. In the case of default, the lender gets the collateral instead of the promised amount. Therefore, the actual delivery of a contract is given by $\min(\varphi, y)$. Agents may promise any non-negative amount φ, so that the set of borrowing contracts is $\mathbb{R}_+$. Note that any contract $\varphi \leq \underline{c}$ is riskless, since the promise is smaller than the minimal asset payoff and will be delivered with certainty. Any contract $\varphi > \underline{c}$ is risky as it may default in some states.

This setting allows agents to chose both the amount borrowed and the promised interest rate. Borrowing an amount of X can be achieved by selling one contract φ_1 with price $q(\varphi_1) = X$, or by selling two identical contracts

[1] Pessimists would like to short-sell the asset, but I do not allow short-selling in this setup.

φ_2 with price $q(\varphi_2) = X/2$. If the pricing function $q(\cdot)$ is concave[2], then $\varphi_1 > 2\varphi_2$. This implies that the promised interest rate in the first case is larger than the one promised in the second case: $\varphi_1/q(\varphi_1) > \varphi_2/q(\varphi_2)$. However, selling two contracts φ_2 requires pledging two units of the asset as collateral, whereas selling one contract φ_1 requires pledging only one unit of the asset. A portfolio of two contracts with smaller promises is much safer for the lender, as it is backed by twice as much of collateral, which allows for a lower interest rate.[3] Holding more of the asset reduces the interest rate on borrowing, which is one of the key features of the model.

2.2.1 Equilibrium Definition

The equilibrium definition in this model is an extension of the one in Simsek (2013) to the case of a continuum of agents. Let $a_i \in \mathbb{R}_+$ denote agent i's asset demand, and let $c_i \in \mathbb{R}_+$ denote the demand for cash. Each agent also decides on borrowing and lending positions on the contract space $\varphi \in \mathbb{R}_+$, denoted by measures μ_i^- and μ_i^+. Measure μ_i^- corresponds to sold contracts, i.e. those used to borrow, whereas measure μ_i^+ – to bought contracts, i.e. those used to lend. Note that it is important to distinguish between contract purchases and contract sells as the latter requires pledging the asset as collateral, whereas the former does not.

Denote agent's endowment by $n := 1 + p$. Then the budget constraint is given by:

$$\underbrace{pa_i}_{\text{Asset Purchases}} + \overbrace{c_i}^{\text{Cash Purchases}} + \underbrace{\int_{\varphi=0}^{\infty} q(\varphi)d\mu_i^+}_{\text{Contract Purchases}} \leq \underbrace{n}_{\text{Endowment}} + \overbrace{\int_{\varphi=0}^{\infty} q(\varphi)d\mu_i^-}^{\text{Contract Sellings}} \tag{BC}$$

where the integrals correspond to total amounts lent and borrowed. For each contract sold, agents need to pledge one unit of the asset as collateral. Therefore, borrowing is limited by the collateral constraint:

$$\underbrace{\int_{\varphi=0}^{\infty} d\mu_i^-}_{\#\text{ Contracts Sold}} \leq \underbrace{a_i}_{\text{Asset Holdings}} \tag{CC}$$

[2]Which is the case in the numerical example in this paper.

[3]Note that in this setting each contract requires the same amount of collateral, independent of the amount promised. It is the contract price that is set in equilibrium to ensure equality of demand and supply. However, one could instead fix the amount borrowed (the price) or the promise for each contract, without affecting the results.

Note that the constraint is not on the amount borrowed, but on the 'number' of contracts sold. The agent's problem is then:

$$\max_{a_i,c_i,\mu_i^+,\mu_i^-} a_i\mathbb{E}_i[y] + c_i + \int_{\varphi=0}^{\infty} \left(\mathbb{E}_i[\min(\varphi,y)]d\mu_i^+ - \mathbb{E}_i[\min(\varphi,y)]d\mu_i^-\right) \quad \text{(AP)}$$
$$\text{subject to (BC) and (CC)}$$

Since agents have linear utility, they maximize the expected payoff of their portfolio. The first term is the expected payoff of asset holdings, and the second term is the payoff of cash holdings. The third term are the expected payoffs of all the contracts agent i bought, and the fourth term are the expected deliveries of all the contracts agent i sold. The market clearing conditions are:

- Asset market: $\int_0^1 a_i di = 1$
- Cash market: $\int_0^1 c_i di = 1$
- Contract markets: $\mu^+ = \int_0^1 \mu_i^+ di = \int_0^1 \mu_i^- di = \mu^-$

The first two conditions are straightforward and state that total asset and cash holdings are equal to total endowments of these commodities. The last condition means that the total measure of contract purchases ($\mu^+ = \int_0^1 \mu_i^+ di$) is equal to the total measure of contract sells ($\mu^- = \int_0^1 \mu_i^- di$).

A general equilibrium is defined as a collection of prices $\{p, q : R_+ \to R_+\}$ and portfolios $\{a_i, c_i, \mu_i^+, \mu_i^-\}_{i\in[0,1]}$ such that portfolios solve (AP) for each $i \in [0, 1]$ and all markets clear.

This formulation of equilibrium is not particularly tractable. There is a continuum of agents making portfolio decisions over a continuum of contracts. In the following I establish several equilibrium properties, which provide insights into its structure and substantially simplify its characterization.

2.3 Equilibrium Properties

Because of linear utility, agents chose a portfolio that maximizes the expected return. There are four positions that each agent may take: buying cash, buying the asset, buying a contract, and buying the asset on margin (buying the asset and selling a contract). Each agent invests her entire endowment into the position that provides the largest (agent-specific) expected return. The return on cash is trivially equal to 1. The return on the asset if purchased unleveraged is given by:

$$\frac{\mathbb{E}_i[y]}{p}$$

The numerator is the expected asset payoff and the denominator is the asset price. If the asset is purchased leveraging with a contract φ (i.e. by buying the asset and selling the contract), the return is denoted by $R_y(i,\varphi)$ and is given by:

$$R_y(i,\varphi) := \frac{\mathbb{E}_i[\max(y-\varphi,0)]}{p-q(\varphi)}$$

When buying the asset leveraged with a contract φ, an agent holds one unit of the asset and promises to repay φ in $t=1$. In case the agent delivers the promise, her payoff is $y-\varphi$. In the case of default, the agent surrenders the asset to the lender and receives 0. Thus the expected payoff of this position is $\mathbb{E}_i[\max(y-\varphi,0)]$. The price is $p-q(\varphi)$, also referred to as the down payment, since part of the asset is financed by borrowing $q(\varphi)$. Finally, the return on buying a contract φ (lending) is denoted by $R_c(i,\varphi)$ and is given by:

$$R_c(i,\varphi) := \frac{\mathbb{E}_i[\min(y,\varphi)]}{q(\varphi)}$$

Given the asset price p and the contract pricing function $q(\cdot)$, each agent computes the highest expected return for each action:

$$\max\left\{1, \frac{\mathbb{E}_i[y]}{p}, \max_{\varphi} R_y(i,\varphi), \max_{\varphi} R_c(i,\varphi)\right\}$$

and takes one position that provides the highest expected return.[4] In fact, only three out of four possible positions can be considered in equilibrium.

Lemma 1.. *In any equilibrium, each agent either buys cash, buys a contract, or buys the asset on margin.*

I provide all the proofs in Appendix B.2. Intuitively, if an agent prefers to buy the asset unleveraged over cash or any contract, she will prefer to leverage on the asset in order to further increase the expected return. I now make an additional assumption on beliefs in order to characterize the equilibrium structure.

Assumption A1.. *Beliefs* $\{F_i(\cdot)\}_{i\in[0,1]}$ *satisfy hazard-rate order:* $\forall i,j \in [0,1]$ *s.t.* $i<j$*:*

$$\frac{f_j(y)}{1-F_j(y)} < \frac{f_i(y)}{1-F_i(y)} \quad \forall y \in (\underline{c},\overline{c})$$

or, equivalently, $\frac{1-F_j(y)}{1-F_i(y)}$ *is strictly increasing in* y *over* $[\underline{c},\overline{c}]$*.*

[4] Or is indifferent between several positions that provide the same expected return and is free to take any combination of these positions.

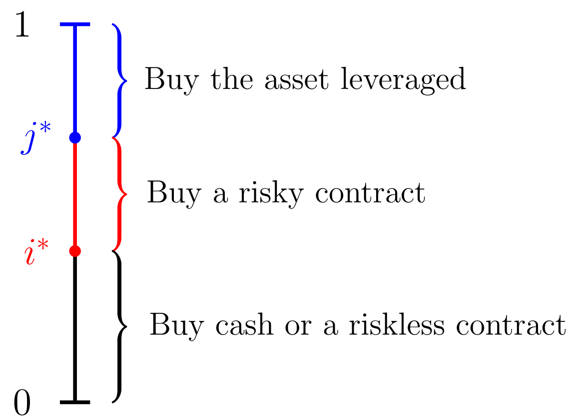
The hazard-rate order is also assumed in Simsek (2013). It implies first order stochastic dominance, but it is weaker than monotone likelihood ratio property. To get an intuition, consider the alternative definition of the hazard-rate order, stating that $\frac{1-F_j(y)}{1-F_i(y)}$ is strictly increasing in y if $j > i$. The numerator $1 - F_j(y)$ is the probability that the asset payoff is larger than some threshold y, assigned by agent j. The fraction $\frac{1-F_j(y)}{1-F_i(y)}$ captures the optimism of agent j relative to the optimism of agent $i < j$, regarding the events 'better' than y. Under hazard-rate order, the ratio is increasing in y, so that j becomes increasingly more optimistic than i as threshold y increases. Assumption A1 disciplines the belief structure and allows for a tighter equilibrium characterization, stated in Theorem 1.

Theorem 1.. *Under Assumption A1, in any equilibrium there exist i^*, j^* such that $j^* > i^*$ and:*

- *Most pessimistic agents ($i < i^*$) buy cash or riskless contracts ($\varphi \leq \underline{c}$)*
- *Most optimistic agents ($j > j^*$) buy the asset leveraged*
- *Agents with moderate optimism ($i^* < i < j^*$) buy risky contracts ($\varphi > \underline{c}$)*

Theorem 1 establishes equilibrium structure, which is illustrated in Figure 2.3.1.

Figure 2.3.1: Equilibrium Structure



Agents sort into three groups. Pessimists prefer safe assets – cash or riskless contracts. Optimists buy the risky asset on margin, and agents with moderate optimism buy risky contracts. The three groups are separated by marginal agents i^* and j^*, so that i^* is indifferent between buying cash and lending, and j^* is indifferent between lending and buying the asset leveraged.

An important part of Theorem 1 is that in any equilibrium there are risky contracts traded, so that any equilibrium features the possibility of default. I now make an assumption on equilibrium contract price function in order to obtain an even sharper equilibrium characterization.

Assumption A2.. *The equilibrium contract price function $q(\cdot)$ is differentiable.*

It is reasonable to expect this property if beliefs $F_i(\cdot)$ are a smooth function of i, so that agents' valuation of any claim on the asset varies gradually with their type. The assumption ensures that contract choices are determined by first-order conditions. If agent i buys a risky contract φ, then it must satisfy $\frac{\partial R_c(i,\varphi)}{\partial \varphi} = 0$. Similarly, if agent j buys the asset leveraging with contract φ, then it satisfies $\frac{\partial R_y(j,\varphi)}{\partial \varphi} = 0$. In addition, Assumption A2 leads to an interesting result, provided in Theorem 2.

Theorem 2.. *Under assumptions A1 and A2, only risky contracts are traded in equilibrium.*

This result is in stark contrast to the one obtained by Fostel and Geanakoplos (2015) for binomial economies, where the state space of the asset payoff consists of two values $\{y_1, y_2\}$. In their model, any equilibrium is equivalent to the one with no default, i.e. the one where only the riskless contract is traded. Their result is overturned in my version of the model, where the space of the asset payoff is a continuum: now only risky contracts are traded in equilibrium, and default is always possible.

However, a richer state space alone does not guarantee the result of Theorem 2. Simsek (2013) considers a model with a continuous asset payoff, but with two agent types only – optimists and pessimists. In his model, riskless contracts are ruled out by assuming that optimists' endowments are not large enough to purchase the risky asset by leveraging with riskless contracts. Relaxing this assumption would lead to the possibility that the riskless contract is the only contract traded in equilibrium. In my setting, endowments and identities of contract buyers and sellers are endogenous. It is thus a combination of a continuous state space and a continuum of agent types that leads to the absence of riskless contracts in equilibrium.

Theorem 3 provides further insights into the equilibrium structure, establishing a one-to-one mapping between contract sellers and contract buyers.

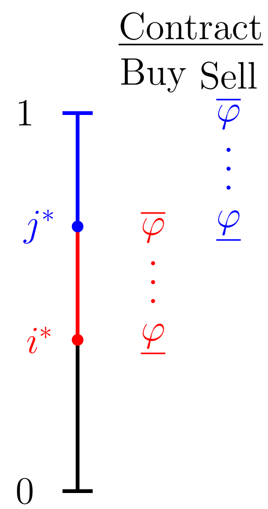
Theorem 3.. *Denote the set of traded contracts by Φ. Under assumptions A1 and A2, in any equilibrium:*

- *Identities of contract buyers are given by a one-to-one strictly increasing function* $i(\varphi) : \Phi \to [i^*, j^*]$
- *Identities of contract sellers are given by a one-to-one strictly increasing function* $j(\varphi) : \Phi \to [j^*, 1]$

Theorem 3 implies that each traded contract φ is associated with a unique pair of agents $\{i(\varphi), j(\varphi)\}$, where $i(\varphi)$ is the buyer and $j(\varphi)$ is the seller. Furthermore, both $i(\varphi)$ and $j(\varphi)$ are increasing in φ, so that riskier contracts are traded by more optimistic agents. Finally, both $i(\varphi)$ and $j(\varphi)$ are bijections, so that each seller and each buyer chose only one risky contract.

The most pessimistic lender i^* is indifferent between buying cash and lending with the safest contract $\underline{\varphi} = \min(\Phi)$. The most optimistic lender j^* is indifferent between lending with the riskiest contract $\overline{\varphi} = \max(\Phi)$ and purchasing the asset on margin by borrowing the safest contract $\underline{\varphi}$. All other lenders and borrowers chose a contract strictly in between the tow bounds. Figure 2.3.2 illustrates the equilibrium contract choices.

Figure 2.3.2: Equilibrium contract choices



I make the final assumption in order to characterize the set of traded contracts Φ.

Assumption A3.. *Distribution function* $F_i(\cdot)$ *is differentiable with respect to agent type* i.

Assumption A3 implies that beliefs vary smoothly across agents. It leads to Lemma 2, which states that every contract between $\underline{\varphi}$ and $\overline{\varphi}$ is traded in equilibrium.

Lemma 2.. *Under assumptions A1:A3, in any equilibrium, the set of traded contracts Φ is an interval $[\underline{\varphi}, \overline{\varphi}]$, with $\underline{\varphi} > \underline{c}$ and $\overline{\varphi} < \overline{c}$.*

By disciplining the belief structure (Assumptions A1, A3) and imposing a restriction on the pricing function (Assumption A2), one obtains a sharp equilibrium characterization. The assumption on the pricing function is clearly strong, as it is imposed on an equilibrium object. However, I find it reasonable, given the amount of details and insights it provides. I now discuss an alternative (tractable) equilibrium definition, using the results from this section.

2.4 Alternative Equilibrium Definition

In the following I revisit the equilibrium definition and propose a more tractable formulation, based on the results from the previous section. The main takeaway is that agents sort into three categories, separated by two marginal agents i^* and j^*. Every agent $i < i^*$ buys cash, any agent $i > j^*$ buys the asset leveraged, and any agent i in between buys a risky contract. Each contract buyer and each contract seller trade only one contract. This implies that the measure of purchased contracts μ_i^+ is a Dirac measure for every contract buyer ($i : i^* < i < j^*$). Similarly, the measure of sold contracts μ_i^- is a Dirac measure for every contract seller ($i > j^*$). Note that $\mu_i^+ = 0$ for all $i < i^*$ and $i > j^*$, whereas $\mu_i^- = 0$ for all $i < j^*$. As a result, μ_i^+ can be summarized by two real numbers: $\{m_i^+, \varphi_i\}$, where φ_i denotes the contract purchased and m_i^+ – the amount. Analogously, μ_i^- can be summarized by $\{m_i^-, \varphi_i\}$. Agent i's portfolio is then given by $x^i = \{c_i, m_i^+, a_i, m_i^-, \varphi_i\}$, where c_i denotes cash holdings and a_i – asset holdings. Note that φ_i corresponds to either bought or sold contracts, depending on whether $m_i^+ > 0$ or $m_i^- > 0$. If agent i buys cash, the value of φ_i is irrelevant.

I make an additional assumption on equilibrium price function $q(\varphi)$ in order to ensure differentiability of lenders' and borrowers' identities $i(\varphi)$ and $j(\varphi)$:

Assumption A2a.. *Equilibrium contract price function $q(\cdot)$ is <u>twice</u> differentiable.*

2.4.1 Agents problem

Denote agent's endowment as before by $n = 1 + p$. Each agent i considers three options:

1. Buying cash. This provides return of 1 for all agents.

2. Buying a contract. In this case, the agent choses which contract to buy (φ_i) and how much of it to buy (m_i^+), maximizing expected payoff $m_i^+\mathbb{E}_i[\min(\varphi_i, y)]$ subject to budget constraint $m_i^+ q(\varphi_i) = n$. Plugging the constraint into the objective yields that the agent is maximizing expected return, choosing from the menu of contracts:

$$\max_{\varphi_i} n\frac{\mathbb{E}_i[\min(\varphi_i, y)]}{q(\varphi_i)} = \max_{\varphi_i} nR_c(i, \varphi_i)$$

The choice is determined by the first order condition:[5]

$$\frac{1 - F_i(\varphi_i)}{R_c(i, \varphi_i)} = q'(\varphi_i) \tag{OB}$$

Denote the optimal lending contract by $\varphi_i^+ = \arg\max_{\varphi_i} R_c(i, \varphi_i)$.

3. Buy the asset leveraged. In this case, the agent choses how much asset to buy (a_i) and which contract to sell (φ_i), using all borrowed funds for further asset purchases, so that the amount of contracts sold m_i^- is equal to the amount of asset bought (a_i). The agent maximizes expected payoff $a_i\mathbb{E}_i[\max(y - \varphi_i, 0)]$ subject to budget constraint $a_i p = n + m_i^- q(\varphi_i)$. Plugging the constraint into the objective yields that the agent is maximizing the expected return, choosing from the menu of contracts:

$$\max_{\varphi_i} n\frac{\mathbb{E}_i[\max(y - \varphi_i, 0)]}{p - q(\varphi_i)} = \max_{\varphi_i} nR_y(i, \varphi_i)$$

The choice is determined by the first order condition:

$$\frac{1 - F_i(\varphi_i)}{R_y(i, \varphi_i)} = q'(\varphi_i) \tag{OS}$$

Denote the optimal borrowing contract by $\varphi_i^- = \arg\max_{\varphi_i} R_y(i, \varphi_i)$.

The three options yield expected returns $\{1, R_c(i, \varphi_i^+), R_y(i, \varphi_i^-)\}$, respectively. Each agent i then choses the option that provides the largest expected return. If agent i prefers to buy cash, then $x^i = \{n, 0, 0, 0, 0\}$. If agent i prefers to buy a contract, then $x^i = \{0, \frac{n}{q(\varphi_i^+)}, 0, 0, \varphi_i^+\}$. If agent i prefers to buy the asset leveraged, then $x^i = \{0, 0, \frac{n}{p - q(\varphi_i^-)}, \frac{n}{p - q(\varphi_i^-)}, \varphi_i^-\}$. Finally, marginal agent i^* is indifferent between buying cash and lending:

$$R_c(i^*, \varphi_{i^*}^+) = 1$$

And marginal agent j^* is indifferent between lending and buying the asset leveraged:

$$R_c(j^*, \varphi_{j^*}^+) = R_y(j^*, \varphi_{j^*}^-)$$

[5] See Apeendix B.1 for derivations.

2.4.2 Market clearing

It is convenient to express market clearing conditions in terms of contracts and identities of buyers and sellers. The set of traded contracts is an interval $[\underline{\varphi}, \overline{\varphi}]$, and identities are given by functions $i(\varphi)$ for the buyers and $j(\varphi)$ for the sellers. These functions are implicitly defined by $\varphi = \arg\max_{\tilde{\varphi}} R_c(i(\varphi), \tilde{\varphi})$ and $\varphi = \arg\max_{\tilde{\varphi}} R_y(j(\varphi), \tilde{\varphi})$, respectively.[6] The first equation states that φ is the optimal lending contract of $i(\varphi)$, and the second equation states that φ is the optimal borrowing contract of $j(\varphi)$. Both functions $i(\varphi)$ and $j(\varphi)$ are differentiable, provided that price function $q(\varphi)$ is twice differentiable. Marginal agents can then be expressed as: $i^* = i(\underline{\varphi})$ and $j^* = i(\overline{\varphi}) = j(\underline{\varphi})$.

The market clearing condition for cash can then be written as:

$$\int_0^1 c_i di = \int_0^{i^*} n di = (1+p)i^* = 1$$

Rewriting this equation gives that:

$$p = \frac{1-i^*}{i^*} = \frac{1-i(\underline{\varphi})}{i(\underline{\varphi})}$$

Recall that agents below i^* do not participate in the market and buy cash, whereas those above i^* either buy the asset or lend to those buying the asset. Thus all the agents above i^* collectively buy the asset. The asset price is then given by the relative size of these two groups – those participating in the market $(1-i^*)$ and those buying cash (i^*).

The asset market clearing condition can be written as:

$$\int_0^1 a_j dj = \int_{j^*}^1 \frac{n}{p - q(\varphi_j^-)} dj = 1$$

Let $\Gamma(\varphi)$ be the share of the asset that is bought by leveraging with contracts $\tilde{\varphi} \leq \varphi$. Then:

$$\Gamma(\varphi) = \int_{j(\underline{\varphi})}^{j(\varphi)} \frac{n}{p - q(\varphi_j^-)} dj = \int_{\underline{\varphi}}^{\varphi} \frac{n}{p - q(\tilde{\varphi})} j'(\tilde{\varphi}) d\tilde{\varphi}$$

where the second equality follows from variable substitution $j(\varphi) \to \varphi$. $\Gamma(\varphi)$ has density $\gamma(\varphi) = \frac{n}{p-q(\varphi)} j'(\varphi)$ and the asset market clearing condition can be written as:

$$\Gamma(\overline{\varphi}) = 1$$

[6] One can also express $i(\varphi)$ and $j(\varphi)$ as inverses of φ_i^+ and φ_j^-, respectively. In that case $i(\varphi_i^+) = i$ and $j(\varphi_j^-) = j$.

In addition, from the definition of $\Gamma(\varphi)$ and $\gamma(\varphi)$ it follows that:

$$\begin{aligned} p\Gamma(\varphi) - \int_{\underline{\varphi}}^{\varphi} q(\tilde{\varphi})\gamma(\tilde{\varphi})d\tilde{\varphi} &= \int_{\underline{\varphi}}^{\varphi} (p - q(\tilde{\varphi}))\gamma(\tilde{\varphi})d\tilde{\varphi} \\ &= n\int_{\underline{\varphi}}^{\varphi} j'(\tilde{\varphi})d\tilde{\varphi} = n\int_{j(\underline{\varphi})}^{j(\varphi)} dj = n\left[j(\varphi) - j(\underline{\varphi})\right] \ \forall\varphi \in [\underline{\varphi}, \overline{\varphi}] \end{aligned}$$

and thus:

$$\underbrace{n\left[j(\varphi) - j(\underline{\varphi})\right]}_{\text{Endowment}} + \underbrace{\int_{\underline{\varphi}}^{\varphi} q(\tilde{\varphi})\gamma(\tilde{\varphi})d\tilde{\varphi}}_{\text{Borrowed}} \quad = \quad \underbrace{p\Gamma(\varphi)}_{\text{Asset Purchases}} \qquad \forall\varphi \in [\underline{\varphi}, \overline{\varphi}] \quad (\text{ABC}.\varphi)$$

I will refer to this condition as the aggregated borrowers budget constraint. The first term on the left-hand side is the total endowment of borrowers that leverage with contracts smaller than φ. The second term is their total borrowings: since each borrowing contract is collateralized by one unit of the asset, $\gamma(\varphi)$ captures the 'amount' of contracts φ sold, and $q(\varphi)$ provides the amount borrowed on each of these contracts. The two terms on the left-hand side give the total amount spent on the asset by agents that sell contracts up to φ. The right-hand side is the price of the asset share, purchased by these agents.

Now consider market clearing condition for contracts:

$$\mu^+ = \int_0^1 \mu_i^+ di = \int_{i^*}^{j^*} \mu_i^+ di \overset{!}{=} \int_{j^*}^1 \mu_j^- dj = \int_0^1 \mu_j^- dj = \mu^-$$

where μ^+ and μ^- are total measures of contracts bought and sold. Since both measures are finite, their equality can be expressed in the following way:

$$\int_{\underline{\varphi}}^{\varphi} d\mu^+ = \int_{\underline{\varphi}}^{\varphi} d\mu^- \ \forall\varphi \in [\underline{\varphi}, \overline{\varphi}]$$

This implies that the two measures coincide on any open interval and thus are equal.[7] Note that $\int_{\underline{\varphi}}^{\varphi} d\mu^-$ is the total amount of contracts $\tilde{\varphi}$ sold, such that $\tilde{\varphi} \in [\underline{\varphi}, \varphi]$. These contracts are sold by agents $j \in [j(\underline{\varphi}), j(\varphi)]$. Agent $j(\tilde{\varphi})$ sells contract $\tilde{\varphi}$ in amount $m_j^-(\tilde{\varphi}) = \frac{n}{p-q(\tilde{\varphi})}$. Thus:

$$\int_{\underline{\varphi}}^{\varphi} d\mu^- = \int_{j(\underline{\varphi})}^{j(\varphi)} \frac{n}{p - q(\varphi_j^-)} dj = \Gamma(\varphi)$$

[7]This can be shown using the Dynkin's $\pi - \lambda$ theorem, with π-system being the set of all open intervals in $\mathbb{R}$ and λ-system defined as $\{A \in \mathcal{B}(\mathbb{R}) : \mu^+(A) = \mu^-(A)\}$.

Analogously, $\int_{\underline{\varphi}}^{\varphi} d\mu^+$ is the total amount of contracts $\tilde{\varphi}$ bought, such that $\tilde{\varphi} \in [\underline{\varphi}, \varphi]$. These contracts are bought by agents $i \in [i(\underline{\varphi}), i(\varphi)]$. Agent $i(\tilde{\varphi})$ buys contract $\tilde{\varphi}$ in amount $m_i^+(\tilde{\varphi}) = \frac{n}{q(\tilde{\varphi})}$. Thus:

$$\int_{\underline{\varphi}}^{\varphi} d\mu^+ = \int_{i(\underline{\varphi})}^{i(\varphi)} \frac{n}{q(\varphi_i^+)} di = \int_{\underline{\varphi}}^{\varphi} \frac{n}{q(\tilde{\varphi})} i'(\tilde{\varphi}) d\tilde{\varphi}$$

and since it must hold for any $\varphi \in [\underline{\varphi}, \overline{\varphi}]$, the contract market clearing condition implies $\gamma(\varphi) = \frac{n}{q(\varphi)} i'(\varphi)$ and thus:

$$\int_{\underline{\varphi}}^{\varphi} q(\tilde{\varphi})\gamma(\tilde{\varphi})d\tilde{\varphi} = n \int_{\underline{\varphi}}^{\varphi} i'(\tilde{\varphi})d\tilde{\varphi} = n \int_{i(\underline{\varphi})}^{i(\varphi)} di = n \left[i(\varphi) - i(\underline{\varphi})\right] \ \forall \varphi \in [\underline{\varphi}, \overline{\varphi}] \Longrightarrow$$

$$n \left[i(\varphi) - i(\underline{\varphi})\right] = \int_{\underline{\varphi}}^{\varphi} q(\tilde{\varphi})\gamma(\tilde{\varphi})d\tilde{\varphi} \quad \forall \varphi \in [\underline{\varphi}, \overline{\varphi}]$$

The term on the right is again the total amount borrowed with contracts up to φ, it appears in the aggregated borrowers budget constraint (ABC.φ). The term on the left is the total endowment of agents that buy contracts up to φ, and since they invest all of their endowment into contracts, it is the total amount lent with contracts up to φ. Contract market clearing requires that the left-hand side is equal to the right-hand side.

An alternative way of writing the contract market clearing condition is via $\gamma(\varphi)$ directly. Using its definition and the market clearing condition yields:

$$\frac{n}{p - q(\varphi)} j'(\varphi) = \frac{n}{q(\varphi)} i'(\varphi)$$

The term on the left is, loosely speaking, the total amount borrowed with contract φ and the term on the right is the total amount lent, so the whole expression is the market clearing condition for contract φ. Note that, intuitively, one would expect this condition to be $\frac{n}{p-q(\varphi)} = \frac{n}{q(\varphi)}$, since $\frac{n}{p-q(\varphi)}$ is how much agent $j(\varphi)$ borrows and $\frac{n}{q(\varphi)}$ is how much agent $i(\varphi)$ lends. However, since both agents are of measure zero, market clearing does not require that these two values coincide. Instead, market clearing accounts for the fact that the mass of agents borrowing with contracts in the vicinity of φ may not be equal to the mass of agents lending with contracts in the vicinity of φ. To see this more clearly, approximate the above expression as follows:

$$\frac{n}{p - q(\varphi)} \frac{j(\varphi + \varepsilon) - j(\varphi - \varepsilon)}{2\varepsilon} = \frac{n}{q(\varphi)} \frac{i(\varphi + \varepsilon) - i(\varphi - \varepsilon)}{2\varepsilon}$$

The leverage of borrowers is given by:

$$\frac{p}{p - q(\varphi)} = 1 + \frac{i(\varphi + \varepsilon) - i(\varphi - \varepsilon)}{j(\varphi + \varepsilon) - j(\varphi - \varepsilon)}$$

and is determined by the relative mass of agents lending and borrowing with contracts close to φ. The larger the mass of lenders is relative to the mass of borrowers, the larger the leverage is.

Recall that the asset price is given by $p = \frac{1-i^*}{i^*}$. Using this result, one can express the aggregate leverage as:

$$\frac{p}{p - \int_{\underline{\varphi}}^{\overline{\varphi}} q(\varphi)\gamma(\varphi)d\varphi} = \frac{p}{p - (1+p)(j^* - i^*)} = \frac{1 - i^*}{1 - j^*} = 1 + \frac{j^* - i^*}{1 - j^*}$$

Thus aggregate leverage is simply the total mass of market participants $(1 - i^*)$ over the total mass of the asset buyers $(1 - j^*)$. Given the expression for the asset price $\left(p = \frac{1-i^*}{i^*}\right)$, one immediately sees that there is no tight link between the asset price and aggregate leverage. The asset price is fully determined by the split of agents into market participants and cash buyers, given by i^*. Leverage, on the contrary, is determined by the split of market participants into lenders and borrowers, given by j^*. If there are more agents taking part in the market, but the proportion of lenders and buyers stays constant, then the asset price increases and aggregate leverage remains constant. If some agents decide to borrow instead of lending, but the total mass of market participants remains constant, then aggregate leverage decreases and the asset price stays the same.

I now have all the necessary ingredients to provide an alternative equilibrium definition, which can then be computed numerically.

2.4.3 Alternative Equilibrium Definition

A general equilibrium consists of:

- A set of traded contracts $[\underline{\varphi}, \overline{\varphi}]$
- Asset price p and contract prices $q(\varphi)$ for $\varphi \geq 0$
- Identities of contract buyers $i(\varphi) \in [i^*, j^*]$ for $\varphi \in [\underline{\varphi}, \overline{\varphi}]$
- Identities of contract sellers $j(\varphi) \in [j^*, 1]$ for $\varphi \in [\underline{\varphi}, \overline{\varphi}]$

Such that:

- Every agent $i < i^*$ optimally buys cash

- Every agent $i : i^* < i < j^*$ optimally buys a contract and the contract choice is given by (OB)
- Every agent $j > j^*$ optimally buys the asset leveraged and the contract choice is given by (OS)
- Marginal agents are indifferent (Ind i^* and Ind j^*)
- Markets clear (MC.1, MC.2, MC.φ)

2.4.4 Solving for the equilibrium

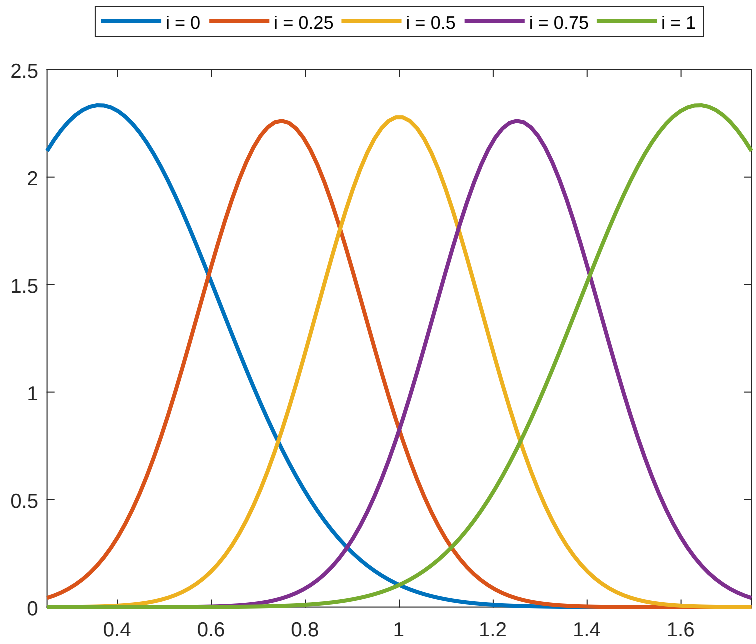
Solving for the equilibrium in this model raises new challenges in comparison with other model variations, considered in the literature. In binomial economies (as in Fostel and Geanakoplos (2015)), the set of contracts traded in equilibrium is directly obtained from the asset payoff structure. In models with two types of agents (as in Simsek (2013), Yan (2017) and Pei and Zhang (2020)), the identities of lenders and borrowers are trivial, and the contract pricing function is pinned down by either indifference or optimality conditions of the lenders. Therefore, in each of these models, certain equilibrium objects can be determined *outside* of equilibrium. In my setting, neither the set of traded contracts, nor the identities of lenders and borrowers, nor the pricing function can be determined without solving for the *entire* equilibrium. I thus resort to numerical methods when solving for the equilibrium.

2.5 Example

In the following I provide a numerical example. Let beliefs $F_i(\cdot)$ be given by a normal distribution, truncated to the interval $[\underline{c}, \overline{c}]$. The parameters of normal distribution (μ_i and σ_i) are calibrated such that for each agent i the expected asset payoff $\mathbb{E}_i[y]$ is equal to $m + (i - 0.5)d$ and the variance of asset payoffs is set to a common value σ^2. Thus, belief structure is parameterized by $\{m, d, \sigma\}$, where m represents average optimism, d – belief disagreement and σ – common uncertainty about the asset payoff.

Let the asset payoff space be given by the interval $[0.25, 1.75]$, average optimism $m = 1$, disagreement $d = 1$, uncertainty $\sigma = 0.175$. Figure 2.5.1 illustrates probability densities for several agents. The equilibrium asset price p for this case is equal to 1.28, the marginal contract buyer i^* is 0.44 and the marginal asset buyer j^* is 0.84. This means that 44% of agents do not participate in the stock market at all and hold cash, 40% lend with risky contracts and only 16% invest in the risky asset. The aggregate leverage can

Figure 2.5.1: Beliefs $f_i(y)$



be computed as $1 + \frac{j^* - i^*}{1 - j^*} = 3.5$. The safest contract traded is $\underline{\varphi} = 0.7$ and the riskiest is $\overline{\varphi} = 1.3$. These values lie well inside the payoff state space $[0.25, 1.75]$. Figure 2.5.2 plots the contract price function $q(\varphi)$ and the credit surface: the promised interest rate $(\varphi/q(\varphi))$ against the Loan-to-Value ratio $(q(\varphi)/p)$.

The contract price function is concave, which means that larger borrowings require increasingly higher promises. The credit surface is convex, which means that higher LTVs require increasingly higher promised interest rates. Recall that the net interest rate on safe cash is 0%. The LTV on the safest contract $\underline{\varphi}$ is equal to 0.54 and is achievable with an interest rate of 1% only. Most optimistic agents, however, finance almost 90% of their asset purchases with debt, which requires promising a 14% interest.

Figure 2.5.3 shows leverage of asset buyers, as well as their shares in the asset. Leverage of the least optimistic buyers is just above 2, whereas the most optimistic agents leverage almost up to the value of 9. Higher leverage allows the most optimistic agents to hold more of the asset and thus the cumulated asset holdings $\Gamma(\varphi)$ is a convex function. In fact, half of the asset is held by the top third of the asset buyers, who account for approximately 5% of the population.

Figure 2.5.2: Contract Prices and Credit Surface

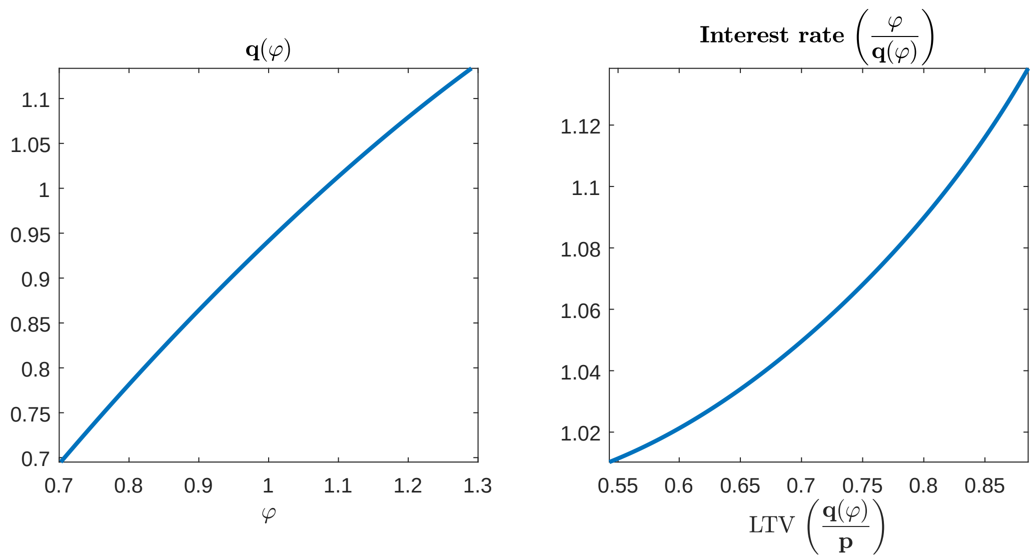
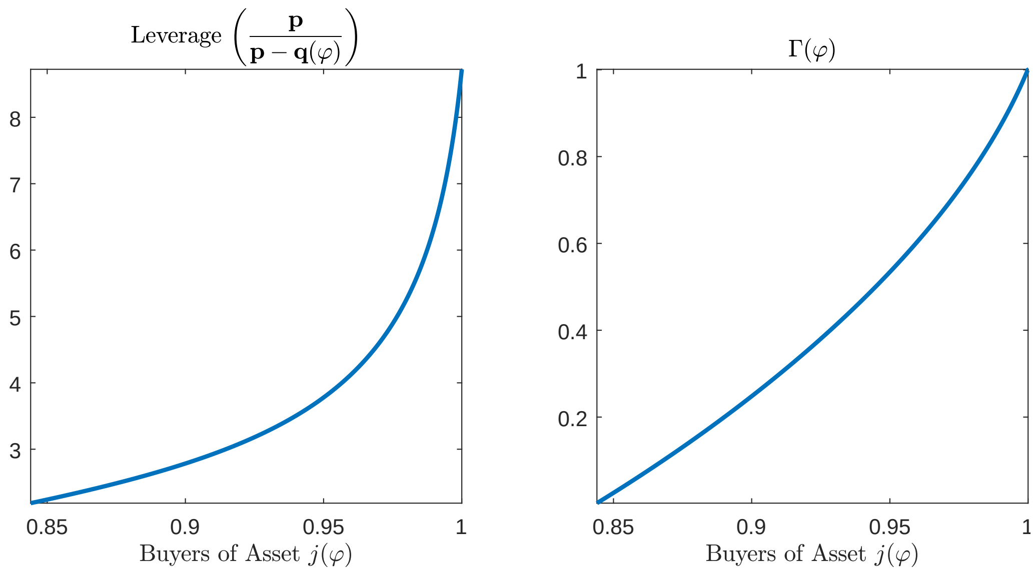


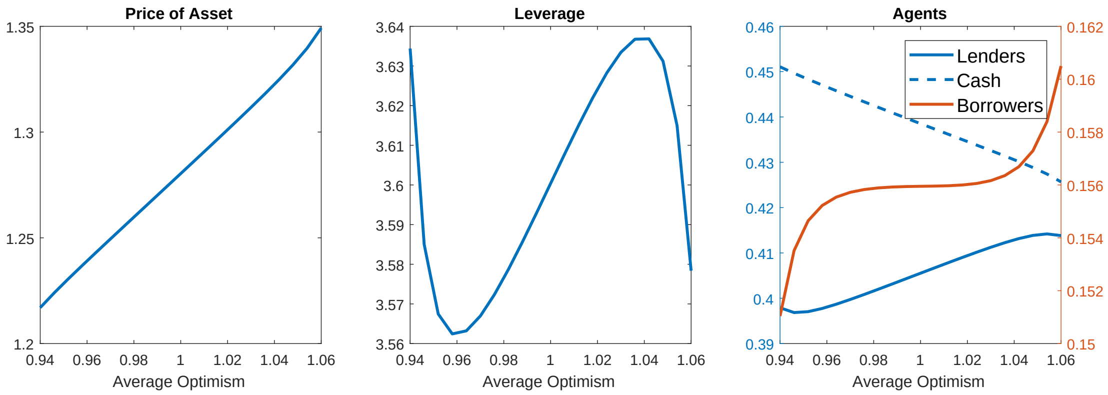
Figure 2.5.3: Buyers of the asset



2.5.1 Varying average optimism

In order to see which primitives affect leverage the most, I first vary average optimism m and study how this changes the equilibrium. Figure 2.5.4 shows the asset price p, aggregate leverage, and the masses of the three groups of agents for different levels of average optimism. The asset price increases linearly with average optimism: the larger the expected value of the asset

Figure 2.5.4: Changing average optimism m



payoff, the more agents want to participate in the market. This is reflected by a steep drop in the share of agents holding cash (the dotted blue line on the third panel, left scale). Leverage, however, is not that sensitive to average optimism, and also does not exhibit a monotone relationship with it. Leverage is determined by the relative mass of lenders and borrowers. Both of these groups grow as optimism increases, but at different rates, which causes small fluctuations in leverage. The mass of asset buyers (the red solid line on the third panel, right scale) increases rapidly for low and high values of optimism, whereas the mass of lenders (the solid blue line on the third panel, left scale) does so for intermediate values of optimism.

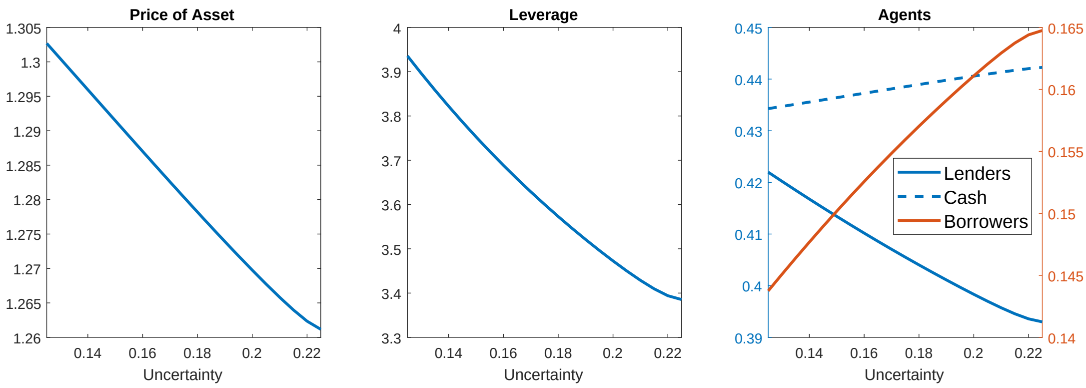
Overall, higher optimism attracts more agents to the market, which increases the price of the asset. The agents that enter the market start lending to the optimists, which increases the pool of lenders. On the other hand, most optimistic lenders decide to start buying the asset instead, which increases the pool of borrowers. Thus the effect on leverage is small and ambiguous.

2.5.2 Varying uncertainty

Figure 2.5.5 provides the same statistics, now for different levels of uncertainty σ. The economy's response to changes in uncertainty is quite different. The asset price is not as sensitive to changes in uncertainty, whereas leverage varies strongly with it.[8] As uncertainty rises, lenders either flee the market

[8] I judge the degree of sensitivity by comparing price and leverage fluctuations across the two exercises. When increasing average optimism from 0.94 to 1.06, I find that the asset price increases by 11%, whereas the magnitude of leverage fluctuations is about 2%. When increasing uncertainty from 0.12 to 0.23, I find that the asset price falls by 3%, whereas leverage falls by 14%. I conclude that price is more sensitive to changes in average optimism, whereas leverage – to changes in uncertainty.

Figure 2.5.5: Changing uncertainty σ



and invest into safe cash, or instead buy the asset and leverage on it. The number of agents willing to lend declines, and the number of agents willing to borrow increases, which leads to lower aggregate leverage. To gain intuition, consider the expected returns of borrowers and lenders. The expected return on lending is given by:

$$R_c(i,\varphi) = \frac{\int_{\underline{c}}^{\varphi} y dF_i(y) + \varphi(1 - F_i(\varphi))}{q(\varphi)}$$

where the first term in the numerator is the expected payoff in the case of default, and the second term is the expected payoff in case the promise is delivered. Note that only the lower part of the distribution ($\underline{c}$ to φ) is relevant for the lender. If uncertainty rises, the distribution spreads out and 'bad' outcomes become 'worse' and more likely, the integral in $R_c(i,\varphi)$ goes down and lowers the expected return for a lender.[9] That is why fewer agents prefer to lend as uncertainty increases. The expected return on buying the asset leveraged is as follows:

$$R_y(i,\varphi) = \frac{\int_{\varphi}^{\bar{c}} y dF_i(y) - \varphi(1 - F_i(\varphi))}{p - q(\varphi)}$$

Here the focus is on the upper part of the distribution (φ to $\underline{c}$). As uncertainty goes up, the 'good' outcomes become even 'better' and more likely, which increases the expected return for a borrower. That is why more agents prefer to buy the asset leveraged as uncertainty rises.

[9] Note that these are partial equilibrium effects, as in general equilibrium the pricing function $q(\varphi)$ will also adjust.

Overall, higher uncertainty scares lenders and forces them to either leave the market or buy the asset instead. This leads to a moderate price decrease and a steep drop in leverage.

2.6 Summary and Avenues for Future Research

In this paper I investigate the determinants of leverage when borrowing is collateralized by an asset, the payoff of which is not affected by asset holders. Typical examples of this setup are REPO loans and mortgages. I show that borrowing in equilibrium is always subject to default and no riskless contracts are traded. Aggregate price and leverage are decoupled: price is determined by the mass of market participants, whereas leverage depends on how they split into lenders and borrowers. A numerical exercise suggests that leverage is very sensitive to uncertainty and decreases sharply when it rises, but does not respond a lot to changes in average optimism.

So far, I have not investigated the effects of belief disagreement on leverage and the asset price, which would be the next natural step in this paper. I expect these effects to be very intricate, as varying belief disagreement has a lot of degrees of freedom. Simsek (2013) finds that it is important whether agents disagree on the probabilities of 'good' or 'bad' states, as these types of disagreement have different implications for the asset price and leverage. My setting allows for an additional dimension of disagreement, namely: which agents become more optimistic and which agents become more pessimistic?

When varying belief disagreement, one has to keep the average optimism fixed in order to identify the pure effects of disagreement. That implies that some agents have to become more optimistic and others more pessimistic. In the numerical example, I set agent i's expected asset payoff to be $m + (i - 0.5)d$, where m is the average optimism and d is the parameter of belief disagreement. One can thus vary d, which would make all agents above the median more optimistic and all agents below the median more pessimistic. However, one could specify more flexible forms of belief disagreement, which would allow to make, e.g. the top 70% of agents more optimistic and the bottom 30% more pessimistic. Then it would matter 'around' which point the disagreement is changing.

Bibliography

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll (2017). *Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach*. Working Paper 23732. National Bureau of Economic Research.

Adam, Klaus and Henning Weber (2019). "Optimal Trend Inflation". In: *American Economic Review* 109.2, pp. 702–37.

— (2020). *Estimating the Optimal Inflation Target from Trends in Relative Prices*. Discussion Paper Series – CRC TR 224 144.

Adrian, Tobias and Hyun Song Shin (2010). "Liquidity and leverage". In: *Journal of Financial Intermediation* 19.3. Risk Transfer Mechanisms and Financial Stability, pp. 418 –437.

Alvarez, Fernando and Francesco Lippi (2014). "Price Setting With Menu Cost for Multiproduct Firms". In: *Econometrica* 82.1, pp. 89–135.

— (2019). *The analytic theory of a monetary shock*. EIEF Working Paper.

Alvarez, Fernando and Andy Neumeyer (Oct. 2019). *The Pass-Through of Large Cost Shocks in an Inflationary Economy*. Working Papers Central Bank of Chile 844. Central Bank of Chile.

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi (2016). "The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach". In: *American Economic Review* 106.10, pp. 2817–51.

Alvarez, Fernando, Martin Beraja, Martín Gonzalez-Rozada, and Pablo Andrés Neumeyer (Feb. 2019). "From Hyperinflation to Stable Prices: Argentina's Evidence on Menu Cost Models". In: *The Quarterly Journal of Economics* 134.1, pp. 451–505.

Alvarez, Fernando, Francesco Lippi, and Aleksei Oskolkov (2020). *The Macroeconomics of Sticky Prices with Generalized Hazard Functions*. Working Paper.

Angrist, Joshua D., Òscar Jordà, and Guido M. Kuersteiner (2018). "Semiparametric Estimates of Monetary Policy Effects: String Theory Revisited". In: *Journal of Business & Economic Statistics* 36.3, pp. 371–387.

Ascari, Guido and Timo Haber (2020). *Non-linearities, state-dependent prices and the transmission mechanism of monetary policy.* Working Paper.

Ascari, Guido and Argia M. Sbordone (2014). “The Macroeconomics of Trend Inflation”. In: *Journal of Economic Literature* 52.3, pp. 679–739.

Auerbach, Alan J. and Yuriy Gorodnichenko (2012). “Measuring the Output Responses to Fiscal Policy”. In: *American Economic Journal: Economic Policy* 4.2, pp. 1–27.

Bachmann, Rüdiger, Ricardo J. Caballero, and Eduardo M. R. A. Engel (2013). “Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model”. In: *American Economic Journal: Macroeconomics* 5.4, pp. 29–67.

Baley, Isaac and Andrés Blanco (2020). *Aggregate Dynamics in Lumpy Economies.* Working Paper.

Ball, Laurence M. (2013). “The Case for Four Percent Inflation”. In: *Central Bank Review* 13.2, pp. 17–31.

Barakchian, S. Mahdi and Christopher Crowe (2013). “Monetary policy matters: Evidence from new shocks data”. In: *Journal of Monetary Economics* 60.8, pp. 950 –966.

Berger, David and Joseph Vavra (2015). “Consumption Dynamics During Recessions”. In: *Econometrica* 83.1, pp. 101–154.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist (1999). “Chapter 21 The financial accelerator in a quantitative business cycle framework”. In: vol. 1. Handbook of Macroeconomics. Elsevier, pp. 1341 –1393.

Blanchard, Olivier, Giovanni Dell’Ariccia, and Paolo Mauro (2010). “Rethinking Macroeconomic Policy”. In: *Journal of Money, Credit and Banking* 42.s1, pp. 199–215.

Blanco, Andrés (2020). “Optimal Inflation Target in an Economy with Menu Costs and a Zero Lower Bound”. In: *American Economic Journal: Macroeconomics (forthcoming).*

Brunnermeier, Markus K. and Lasse Heje Pedersen (June 2009). “Market Liquidity and Funding Liquidity”. In: *The Review of Financial Studies* 22.6, pp. 2201–2238.

Brunnermeier, Markus K. and Yuliy Sannikov (2014). “A Macroeconomic Model with a Financial Sector”. In: *American Economic Review* 104.2, pp. 379–421.

Caballero, Ricardo J. and Eduardo M.R.A. Engel (2007). “Price stickiness in Ss models: New interpretations of old results”. In: *Journal of Monetary Economics* 54, pp. 100 –121.

Calvo, Guillermo A. (1983). “Staggered prices in a utility-maximizing framework”. In: *Journal of Monetary Economics* 12.3, pp. 383 –398.

Cavallo, Alberto (2018). “Scraped Data and Sticky Prices”. In: *The Review of Economics and Statistics* 100.1, pp. 105–119.

Coibion, Olivier (2012). “Are the Effects of Monetary Policy Shocks Big or Small?” In: *American Economic Journal: Macroeconomics* 4.2, pp. 1–32.

Coibion, Olivier, Yuriy Gorodnichenko, and Johannes Wieland (Mar. 2012). “The Optimal Inflation Rate in New Keynesian Models: Should Central Banks Raise Their Inflation Targets in Light of the Zero Lower Bound?” In: *The Review of Economic Studies* 79.4, pp. 1371–1406.

Coimbra, Nuno and Hélène Rey (2020). *Financial Cycles with Heterogeneous Intermediaries*. Working Paper.

Cover, James Peery (Nov. 1992). “Asymmetric Effects of Positive and Negative Money-Supply Shocks*”. In: *The Quarterly Journal of Economics* 107.4, pp. 1261–1282.

Diercks, Anthony M. (2017). *The Reader’s Guide to Optimal Monetary Policy*. Available at SSRN: 2989237.

Fostel, Ana and John Geanakoplos (2008). “Leverage Cycles and the Anxious Economy”. In: *American Economic Review* 98.4, pp. 1211–44.

— (2015). “Leverage and Default in Binomial Economies: A Complete Characterization”. In: *Econometrica* 83.6, pp. 2191–2229.

Geanakoplos, John (1997). “Promises, Promises”. In: *The Economy As An Evolving Complex System II*. Reading, MA: Addison-Wesley.

— (2003). “Liquidity, Default, and Crashes: Endogenous Contracts in General Equilibrium”. In: *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*. Ed. by Mathias Dewatripont, Lars Peter Hansen, and Stephen J.Editors Turnovsky. Vol. 2. Econometric Society Monographs. Cambridge University Press, 170–205.

— (2010). “The Leverage Cycle”. In: *NBER Macroeconomics Annual 2009, Volume 24*. University of Chicago Press, pp. 1–65.

Geanakoplos, John and William R. Zame (2014). “Collateral equilibrium, I: a basic framework”. In: *Economic Theory* 56.3, pp. 443–492.

Gertler, Mark and Peter Karadi (2015). “Monetary Policy Surprises, Credit Costs, and Economic Activity”. In: *American Economic Journal: Macroeconomics* 7.1, pp. 44–76.

Golosov, Mikhail and Robert E. Lucas (2007). “Menu Costs and Phillips Curves”. In: *Journal of Political Economy* 115.2, pp. 171–199.

He, Zhiguo and Arvind Krishnamurthy (2013). “Intermediary Asset Pricing”. In: *American Economic Review* 103.2, pp. 732–70.

Jaimovich, Nir and Henry E. Siu (2020). “Job Polarization and Jobless Recoveries”. In: *The Review of Economics and Statistics* 102.1, pp. 129–147.

Jarociński, Marek and Peter Karadi (2020). "Deconstructing Monetary Policy Surprises—The Role of Information Shocks". In: *American Economic Journal: Macroeconomics* 12.2, pp. 1–43.

Jordà, Òscar (2005). "Estimation and Inference of Impulse Responses by Local Projections". In: *American Economic Review* 95.1, pp. 161–182.

Karadi, Peter and Adam Reiff (2019). "Menu Costs, Aggregate Fluctuations, and Large Shocks". In: *American Economic Journal: Macroeconomics* 11.3, pp. 111–46.

Kehoe, Patrick and Virgiliu Midrigan (2015). "Prices are sticky after all". In: *Journal of Monetary Economics* 75, pp. 35 –53. ISSN: 0304-3932.

Khan, Aubhik and Julia K. Thomas (2008). "Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics". In: *Econometrica* 76.2, pp. 395–436.

Kiyotaki, Nobuhiro and John Moore (1997). "Credit Cycles". In: *Journal of Political Economy* 105.2, pp. 211–248.

L'Huillier, Jean-Paul and Raphael Schoenle (2020). *Raising the Inflation Target:How Much Extra Room Does It Really Give?* Working Paper.

Lo, Ming Chien and Jeremy Piger (2005). "Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model". In: *Journal of Money, Credit and Banking* 37.5, pp. 865–886.

Long, J. Bradford de and Lawrence H. Summers (1988). "How Does Macroeconomic Policy Affect Output?" In: *Brookings Papers on Economic Activity* 1988.2, pp. 433–494.

Mankiw, N. Gregory and Laurence Ball (Feb. 1994). "Asymmetric Price Adjustment and Economic Fluctuations". In: *Economic Journal* 104, pp. 247–61.

Midrigan, Virgiliu (2011). "Menu Costs, Multiproduct Firms, and Aggregate Fluctuations". In: *Econometrica* 79.4, pp. 1139–1180.

Nakamura, Emi and Jón Steinsson (Nov. 2008). "Five Facts about Prices: A Reevaluation of Menu Cost Models". In: *The Quarterly Journal of Economics* 123.4, pp. 1415–1464.

— (Aug. 2010). "Monetary Non-neutrality in a Multisector Menu Cost Model*". In: *The Quarterly Journal of Economics* 125.3, pp. 961–1013.

— (Jan. 2018). "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect". In: *The Quarterly Journal of Economics* 133.3, pp. 1283–1330.

Pei, Guangyu and Tong Zhang (2020). *Haircut Cycles*. Working Paper.

Ramey, Valerie A and Sarah Zubairy (2014). *Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data.* Working Paper 20719. National Bureau of Economic Research.

Sheshinski, Eytan and Yoram Weiss (1977). "Inflation and Costs of Price Adjustment". In: *The Review of Economic Studies* 44.2, pp. 287–303.

Shleifer, Andrei and Robert W. Vishny (1992). "Liquidation Values and Debt Capacity: A Market Equilibrium Approach". In: *The Journal of Finance* 47.4, pp. 1343–1366.

Simsek, Alp (2013). "Belief Disagreements and Collateral Constraints". In: *Econometrica* 81.1, pp. 1–53.

Stokey, Nancy L. (2009). *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. Princeton University Press.

Tenreyro, Silvana and Gregory Thwaites (2016). "Pushing on a String: US Monetary Policy Is Less Powerful in Recessions". In: *American Economic Journal: Macroeconomics* 8.4, pp. 43–74.

Tsiddon, Daniel (Oct. 1993). "The (Mis)Behaviour of the Aggregate Price Level". In: *The Review of Economic Studies* 60.4, pp. 889–902.

Vavra, Joseph (Feb. 2014). " Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation ". In: *The Quarterly Journal of Economics* 129.1, pp. 215–258.

Yan, Ji (2017). "Belief Heterogeneity, Collateral and Endogenous Leverage". Working paper.

Yun, Tack (1996). "Nominal price rigidity, money supply endogeneity, and business cycles". In: *Journal of Monetary Economics* 37.2, pp. 345 –370.

Appendices

Appendix A

Chapter 1

A.1 Miscellaneous results

A.1.1 Expressing $M(\delta, \mu)$ in terms of price gaps

Alvarez and Lippi (2014) show that impulse response of aggregate price level can be approximated as:

$$P(t) - \bar{P}(t) \approx \delta + \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} z dF_t(z, \mu) - \bar{x}(\mu) \quad \text{with} \quad \bar{x}(\mu) = \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} z dF(z, \mu)$$

where $P(t)$ is the aggregate log-price t periods after shock δ, $\bar{P}(t)$ is the hypothetical price in absence of shock, $F_t(z, \mu)$ is the period t distribution of price gaps, $F(z, \mu)$ is the stationary distribution of price gaps and $\bar{x}(\mu)$ is the average price gap in steady state. Note that instead of evolution of gap distribution $(F_t(z, \mu))$, one can consider conditional evolution of gaps given initial after-shock distribution $F_\delta(z, \mu)$:

$$P(t) - \bar{P}(t) \approx \delta + \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \mathbb{E}\left(z(t) - \bar{x}(\mu) \;\middle|\; z(0) = z\right) dF_\delta(z, \mu)$$

where $\bar{x}(\mu)$ is taken inside integral and expectation. Finally, switching the order of integration, one obtains:

$$\begin{aligned} M(\delta, \mu) &= \int_0^\infty \left[\delta - \left(P(t) - \bar{P}(t)\right)\right] dt \\ &\approx - \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \mathbb{E}\left(\int_0^\infty \left(z(t) - \bar{x}(\mu)\right) dt \;\middle|\; z(0) = z\right) dF_\delta(z, \mu) \end{aligned}$$

A.1.2 Driftless Benchmark

Suppose $\mu = 0$. Firms' value function satisfies the following HJB:

$$\rho v(z) = -z^2 + \frac{\sigma^2}{2} v''(z)$$

The general solution to which is:

$$v(z) = A(e^{\alpha z} + e^{-\alpha z}) - \frac{1}{\rho} z^2 - \frac{\sigma^2}{\rho^2}$$

where $\alpha = \sqrt{2\rho/\sigma^2}$ and A is the unknown coefficient that depends on boundary conditions. These are given by $v(\underline{z}(0)) = v(\overline{z}(0)) = v(\hat{z}(0)) - \kappa$. Due to symmetry, $\underline{z}(0) = -\overline{z}(0)$ and $\hat{z}(0) = 0$. Denote $\overline{z}_0 = \overline{z}(0)$ to ease notation.

Using the expression for $v(z)$ and combining one of the boundary conditions with smooth pasting condition $v'(\overline{z}_0) = 0$, one gets:

$$A = \frac{2\overline{z}_0}{\alpha\rho(e^{\alpha\overline{z}_0} - e^{-\alpha\overline{z}_0})}$$
$$\overline{z}_0^2 = A\rho(e^{\alpha\overline{z}_0} + e^{-\alpha\overline{z}_0} - 2) + \rho\kappa$$

which implicitly defines solution triplet $\{-\overline{z}_0, 0, \overline{z}_0\}$.

Stationary density is defined by Kolmogorov forward equation $(\sigma^2/2)f_{zz}(z,0) = 0$ with boundary conditions $f(\overline{z}_0, 0) = f(-\overline{z}_0, 0) = 0$, integration to one $\int_{-\overline{z}_0}^{\overline{z}_0} f(z,0)dz = 1$ and continuity at $z = 0$. It is thus given by:

$$f(z,0) = \frac{\overline{z}_0 - |z|}{\overline{z}_0^2}$$

for all $z \in [-\overline{z}_0, \overline{z}_0]$ and is zero otherwise. For completeness, cumulative distribution function $F(z,0)$ is then given by:

$$F(z,0) = \begin{cases} \frac{(\overline{z}_0+z)^2}{2\overline{z}_0^2}, & \text{for } z < 0 \\ 1 - \frac{(\overline{z}_0-z)^2}{2\overline{z}_0^2}, & \text{for } z \geq 0 \end{cases}$$

Consider a positive shock $\delta > 0$. Impact effect in driftless economy is:

$$\Theta(\delta, 0) = -\int_{-\overline{z}_0-\delta}^{-\overline{z}_0} zf(z+\delta, 0)dz$$

and due to a kink in $f(z,0)$ is computed separately for smaller ($\delta \leq \overline{z}_0$) and larger ($\delta \geq \overline{z}_0$) shocks. A direct computation of the integral provides:

$$\Theta(\delta,0) = \begin{cases} \frac{1}{6\overline{z}_0^2}\delta^2(\delta + 3\overline{z}_0), & \text{for } \delta < \overline{z}_0 \\ \frac{1}{6\overline{z}_0^2}\left[\delta(6\overline{z}_0^2 + 3\delta\overline{z}_0 - \delta^2) - 4\overline{z}_0^3\right], & \text{for } \delta \in [\overline{z}_0, 2\overline{z}_0) \\ \delta, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}$$

The last line follows since for any $\delta \geq 2\overline{z}_0$:

$$\Theta(\delta,0) = -\int_{-\overline{z}_0-\delta}^{-\overline{z}_0} zf(z+\delta,0)dz = -\int_{-\overline{z}_0}^{-\overline{z}_0+\delta} (z-\delta)f(z,0)dz$$
$$= -\int_{-\overline{z}_0}^{\overline{z}_0} (z-\delta)f(z,0)dz = \delta - \bar{x}(0) = \delta$$

where $\bar{x}(0) = \int_{-\overline{z}_0}^{\overline{z}_0} zf(z,0)dz$ is the average gap. First equality is due to variable substitution, second follows from the fact that $f(z,0) = 0$ for $z \geq \overline{z}_0$ and third one is immediate. Finally, $\bar{x}(0) = 0$ due to symmetry of $f(z,0)$.

Consider now cumulative impulse response ($\delta > 0$):

$$M(\delta, 0) = -\int_{-\overline{z}_0}^{\overline{z}_0} \mathbb{E}\left(\int_0^{\tau} z(t)dt \mid z(0) = z\right) dF_\delta(z,0)$$

Define $m(z,0)$ to be the expected cumulated price gap until first adjustment, so that $m(z,0) = \mathbb{E}\left(\int_0^{\tau} z(t)dt \mid z(0) = z\right)$. The second argument of the function highlights that it is evaluated under $\mu = 0$. This function is characterized by $z + (\sigma^2/2)m_{zz}(z,0) = 0$ together with boundary conditions $m(\overline{z}_0, 0) = m(-\overline{z}_0, 0) = 0$, which implies that:

$$m(z,0) = \frac{\overline{z}_0^2 z - z^3}{3\sigma^2}$$

Given that shock shifts the entire distribution in parallel and some firms adjust immediately, distribution $F_\delta(z,0)$ is the shifted stationary distribution, so that $F_\delta(z,0) = F(z+\delta, 0)$ for all $z \in [\underline{z}, \overline{z}-\delta]$ and $F_\delta(z,0) = 1$ for all $z \in (\overline{z}-\delta, \overline{z}]$. In addition, there is a mass point at $z = 0$ due to firms that adjust immediately, equal to $F(-\overline{z}_0+\delta, 0)$. $M(\delta, 0)$ is then given by:

$$M(\delta, 0) = -\int_{-\overline{z}_0}^{\overline{z}_0-\delta} m(z,0)f(z+\delta,0)dz + m(0,0)F(-\overline{z}_0+\delta, 0)$$

where the second term can be ignored since $m(0,0) = 0$. Again, due to a kink in $f(z,0)$, the integral has to be considered separately for smaller ($\delta \leq \overline{z}_0$) and larger ($\delta \geq \overline{z}_0$) shocks. A direct computation yields:

$$M(\delta, 0) = \begin{cases} \frac{1}{180\sigma^2\overline{z}_0^2}[3\delta^5 + 15\delta^4\overline{z}_0 - 40\delta^3\overline{z}_0^2 + 30\delta\overline{z}_0^4], & \text{for } \delta < \overline{z}_0 \\ \frac{1}{180\sigma^2\overline{z}_0^2}[-3\delta^5 + 15\delta^4\overline{z}_0 - 20\delta^3\overline{z}_0^2 + 16\overline{z}_0^5], & \text{for } \delta \in [\overline{z}_0, 2\overline{z}_0) \\ 0, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}$$

The last line is trivial since if $\delta \geq 2\overline{z}_0$, then the integral in $M(\delta, 0)$ is taken over the interval $[\overline{z}_0-\delta, -\overline{z}_0]$, where $F_\delta(z,0)$ has no mass.

A.1.3 Optimal policy under non-zero drift

Recall that firm's value function solves the following HJB equation for any $z \in [\underline{z}(\mu), \overline{z}(\mu)]$:

$$\rho v(z) = -z^2 - \mu v'(z) + \frac{\sigma^2}{2} v''(z)$$

General solution to $v(z)$ is thus given by:

$$v(z) = C_1 e^{R_1 z} + C_2 e^{R_2 z} - \frac{1}{\rho} z^2 + \frac{2\mu}{\rho^2} z - \left(\frac{\sigma^2}{\rho^2} + \frac{2\mu^2}{\rho^3} \right)$$

where $R_1 = \frac{\mu - \sqrt{\mu^2 + 2\sigma^2 \rho}}{\sigma^2}$, $R_2 = \frac{\mu + \sqrt{\mu^2 + 2\sigma^2 \rho}}{\sigma^2}$. Coefficients C_1 and C_2 are unknown and determined by boundary conditions $v(\underline{z}) = v(\overline{z}) = v(\hat{z}) - \kappa$, where I drop the argument μ in policy variables in order to ease notation. In addition $v(z)$ satisfies smooth pasting conditions $v'(\underline{z}) = v'(\overline{z}) = 0$ and optimality condition $v'(\hat{z}) = 0$. Altogether this results in a system of equations:

$$\begin{aligned}
C_1 R_1 e^{R_1 \underline{z}} + C_2 R_2 e^{R_2 \underline{z}} - \frac{2}{\rho} \underline{z} + \frac{2\mu}{\rho^2} &= 0 (h_1) \\
C_1 R_1 e^{R_1 \hat{z}} + C_2 R_2 e^{R_2 \hat{z}} - \frac{2}{\rho} \hat{z} + \frac{2\mu}{\rho^2} &= 0 \\
C_1 R_1 e^{R_1 \overline{z}} + C_2 R_2 e^{R_2 \overline{z}} - \frac{2}{\rho} \overline{z} + \frac{2\mu}{\rho^2} &= 0 \\
C_1 (e^{R_1 \underline{z}} - e^{R_1 \hat{z}}) + C_2 (e^{R_2 \underline{z}} - e^{R_2 \hat{z}}) - \frac{1}{\rho} (\underline{z}^2 - \hat{z}^2) + \frac{2\mu}{\rho^2} (\underline{z} - \hat{z}) + \kappa &= 0 \\
C_1 (e^{R_1 \overline{z}} - e^{R_1 \hat{z}}) + C_2 (e^{R_2 \overline{z}} - e^{R_2 \hat{z}}) - \frac{1}{\rho} (\overline{z}^2 - \hat{z}^2) + \frac{2\mu}{\rho^2} (\overline{z} - \hat{z}) + \kappa &= 0 (h_5)
\end{aligned}$$

Let ψ denote the vector of unknowns: $\psi = [\underline{z}, \hat{z}, \overline{z}, C_1, C_2]$. Then the above system of equations can be summarized as:

$$H(\mu, \psi) = \mathbf{0} \tag{A.4}$$

where $H : \mathbb{R} \times \mathbb{R}^5 \to \mathbb{R}^5$ and each row of $H(\mu, \psi)$ corresponds to one of the equations (h_1) – (h_5). Given μ, equation (A.4) implicitly defines solution triplet $\{\underline{z}, \hat{z}, \overline{z}\}$ and coefficients C_1 and C_2. Applying Implicit Function Theorem yields:

$$\left. \frac{\partial \psi}{\partial \mu} \right|_{\mu=0} = - \left[\left. \frac{\partial H}{\partial \psi} \right|_{\mu=0} \right]^{-1} \left. \frac{\partial H}{\partial \mu} \right|_{\mu=0}$$

provided $\left. \frac{\partial H}{\partial \psi} \right|_{\mu=0}$ has full rank. Recall from Appendix A.1.2 that under $\mu = 0$ solution to (A.4) is $\psi_0 = [-\overline{z}_0, 0, \overline{z}_0, A, A]$, where $\overline{z}_0$ and A satisfy:

$$\begin{aligned}
A &= \frac{2\overline{z}_0}{\alpha \rho (e^{\alpha \overline{z}_0} - e^{-\alpha \overline{z}_0})} \\
\overline{z}_0^2 &= A\rho (e^{\alpha \overline{z}_0} + e^{-\alpha \overline{z}_0} - 2) + \rho\kappa
\end{aligned}$$

with $\alpha = \sqrt{2\rho/\sigma^2}$. Let $w_1 = (e^{\alpha\bar{z}_0} - e^{-\alpha\bar{z}_0})$, $w_2 = (e^{\alpha\bar{z}_0} + e^{-\alpha\bar{z}_0})$, $\gamma = \frac{2\alpha\bar{z}_0 w_2 - 2w_1}{\rho w_1}$ and $\beta = \frac{4\alpha\bar{z}_0 - 2w_1}{\rho w_1}$. Then a direct computation provides:

$$\left.\frac{\partial H}{\partial \psi}\right|_{\mu=0} = \begin{bmatrix} \gamma & 0 & 0 & -\alpha e^{\alpha\bar{z}_0} & \alpha e^{-\alpha\bar{z}_0} \\ 0 & \beta & 0 & -\alpha & \alpha \\ 0 & 0 & \gamma & -\alpha e^{-\alpha\bar{z}_0} & \alpha e^{\alpha\bar{z}_0} \\ 0 & 0 & 0 & e^{\alpha\bar{z}_0} - 1 & e^{-\alpha\bar{z}_0} - 1 \\ 0 & 0 & 0 & e^{-\alpha\bar{z}_0} - 1 & e^{\alpha\bar{z}_0} - 1 \end{bmatrix}$$

This matrix can be inverted as:

$$\left[\left.\frac{\partial H}{\partial \psi}\right|_{\mu=0}\right]^{-1} = \begin{bmatrix} \gamma^{-1} & 0 & 0 & \frac{\alpha(w_2+1)}{\gamma w_1} & \frac{\alpha}{\gamma w_1} \\ 0 & \beta^{-1} & 0 & \frac{\alpha}{\beta w_1} & -\frac{\alpha}{\beta w_1} \\ 0 & 0 & \gamma^{-1} & -\frac{\alpha}{\gamma w_1} & -\frac{\alpha(w_2+1)}{\gamma w_1} \\ 0 & 0 & 0 & \frac{w_1+w_2-2}{2w_1(w_2-2)} & \frac{w_1-w_2+2}{2w_1(w_2-2)} \\ 0 & 0 & 0 & \frac{w_1-w_2+2}{2w_1(w_2-2)} & \frac{w_1+w_2-2}{2w_1(w_2-2)} \end{bmatrix}$$

The derivative of $H(\mu, \psi)$ with respect to μ evaluated at $\mu = 0$ is:

$$\left.\frac{\partial H}{\partial \mu}\right|_{\mu=0} = \begin{bmatrix} \frac{\alpha\bar{z}_0 w_2 + \alpha^2\bar{z}_0^2 w_1 + 2w_1}{\rho^2 w_1} \\ 2\frac{\alpha\bar{z}_0 + w_1}{\rho^2 w_1} \\ \frac{\alpha\bar{z}_0 w_2 + \alpha^2\bar{z}_0^2 w_1 + 2w_1}{\rho^2 w_1} \\ -\frac{\alpha^2\bar{z}_0^2 w_2 + 2\alpha\bar{z}_0 w_1}{\alpha\rho^2 w_1} \\ \frac{\alpha^2\bar{z}_0^2 w_2 + 2\alpha\bar{z}_0 w_1}{\alpha\rho^2 w_1} \end{bmatrix}$$

Multiplying and collecting terms yields:

$$\left.\frac{\partial \underline{z}}{\partial \mu}\right|_{\mu=0} = \left.\frac{\partial \bar{z}}{\partial \mu}\right|_{\mu=0} = \frac{4\alpha^2\bar{z}_0^2 + \alpha\bar{z}_0 w_1 w_2 - 2w_1^2}{2\rho(\alpha\bar{z}_0 w_1 w_2 - w_1^2)}$$

$$\left.\frac{\partial \hat{z}}{\partial \mu}\right|_{\mu=0} = \frac{\alpha^2\bar{z}_0^2 w_2 + \alpha\bar{z}_0 w_1 - w_1^2}{\rho(2\alpha\bar{z}_0 w_1 - w_1^2)}$$

In order to recover the no-discounting case of Alvarez et al. (2019), use expressions from Appendix A.2.1 and expand numerators up to 6th degree and denominators up to 4th degree.

A.1.4 Stationary density under non-zero drift

Both the impact and cumulative impulse responses depend on stationary density. In this section I provide derivatives of the stationary density function $f(z, \mu)$ with respect to drift μ, evaluated at $\mu = 0$.

Recall that stationary density satisfies the following Kolmogorov forward equation:

$$0 = \mu f_z(z,\mu) + \frac{\sigma^2}{2} f_{zz}(z,\mu)$$

together with boundary conditions $f(\underline{z}(\mu),\mu) = f(\overline{z}(\mu),\mu) = 0$, unit mass condition $\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} f(z,\mu)dz = 1$ and continuity at $z = \hat{z}(\mu)$. Note that density depends on drift μ both directly as it appears in KFE, and indirectly as it also appears in boundary conditions via policy variables. For the purpose of derivation, it is thus convenient to include policy variables explicitly as arguments with some abuse of notation: $f(z;\mu,\underline{z},\hat{z},\overline{z})$, so that $f(z,\mu) = f(z;\mu,\underline{z}(\mu),\hat{z}(\mu),\overline{z}(\mu))$. Stokey (2009) shows that stationary density is given by[1]:

$$f(z;\mu,\underline{z},\hat{z},\overline{z}) = \begin{cases} \frac{e^{\eta(\mu)\hat{z}} - e^{\eta(\mu)\overline{z}} + e^{\eta(\mu)(\overline{z}+\underline{z}-z)} - e^{\eta(\mu)(\hat{z}+\underline{z}-z)}}{(\overline{z}-\underline{z})e^{\eta(\mu)\hat{z}} - (\overline{z}-\hat{z})e^{\eta(\mu)\underline{z}} - (\hat{z}-\underline{z})e^{\eta(\mu)\overline{z}}} & \text{for } z < \hat{z} \\[2ex] \frac{e^{\eta(\mu)\hat{z}} - e^{\eta(\mu)\underline{z}} + e^{\eta(\mu)(\overline{z}+\underline{z}-z)} - e^{\eta(\mu)(\hat{z}+\overline{z}-z)}}{(\overline{z}-\underline{z})e^{\eta(\mu)\hat{z}} - (\overline{z}-\hat{z})e^{\eta(\mu)\underline{z}} - (\hat{z}-\underline{z})e^{\eta(\mu)\overline{z}}} & \text{for } z \geq \hat{z} \end{cases}$$

where $\eta(\mu) = 2\mu/\sigma^2$. To ease notation and simplify later derivations, define $v(z,\mu) = e^{\eta(\mu)z}$, so that:

$$f(z;\mu,\underline{z},\hat{z},\overline{z}) = \begin{cases} \frac{v(\hat{z},\mu) - v(\overline{z},\mu) + v(\overline{z}+\underline{z}-z,\mu) - v(\hat{z}+\underline{z}-z,\mu)}{(\overline{z}-\underline{z})v(\hat{z},\mu) - (\overline{z}-\hat{z})v(\underline{z},\mu) - (\hat{z}-\underline{z})v(\overline{z},\mu)} & \text{for } z < \hat{z} \\[2ex] \frac{v(\hat{z},\mu) - v(\underline{z},\mu) + v(\overline{z}+\underline{z}-z,\mu) - v(\hat{z}+\overline{z}-z,\mu)}{(\overline{z}-\underline{z})v(\hat{z},\mu) - (\overline{z}-\hat{z})v(\underline{z},\mu) - (\hat{z}-\underline{z})v(\overline{z},\mu)} & \text{for } z \geq \hat{z} \end{cases}$$

Partial derivative of $f(z,\mu)$ with respect to drift is the total derivative of $f(z,\mu,\underline{z},\hat{z},\overline{z})$:

$$\begin{aligned} \left.\frac{\partial f(z,\mu)}{\partial \mu}\right|_{\mu=0} &= \left.\frac{df(z;\mu,\underline{z}(\mu),\hat{z}(\mu),\overline{z}(\mu))}{d\mu}\right|_{\mu=0,\underline{z}=\underline{z}(0),\hat{z}=\hat{z}(0),\overline{z}=\overline{z}(0)} \\ &= \left.\left(\frac{\partial f(z;\cdot)}{\partial \mu} + \frac{\partial f(z;\cdot)}{\partial \underline{z}}\frac{\partial \underline{z}}{\partial \mu} + \frac{\partial f(z;\cdot)}{\partial \hat{z}}\frac{\partial \hat{z}}{\partial \mu} + \frac{\partial f(z;\cdot)}{\partial \overline{z}}\frac{\partial \overline{z}}{\partial \mu}\right)\right|_{\mu=0,\underline{z}=\underline{z}(0),\hat{z}=\hat{z}(0),\overline{z}=\overline{z}(0)} \end{aligned}$$

the first component is the direct effect of μ on the shape of stationary density, whereas the latter three are indirect effects through optimal policy. Recall that $\hat{z}(0) = 0$ and note that due to symmetry of density around $\mu = 0$, $f(-z,\mu) = f(z,-\mu)$ and thus $\frac{df(-z,0)}{d\mu} = -\frac{df(z,0)}{d\mu}$, so it suffices to calculate the derivative for $z < 0$ only.

[1] See Chapter 5. The formula is obtained as $f(z) = L(z)/\tau$, where $L(z)$ is the expected local time at z and τ is the average length between adjustments.

The three derivatives $\left\{\frac{\partial f(z;\cdot)}{\partial \underline{z}}, \frac{\partial f(z;\cdot)}{\partial \hat{z}}, \frac{\partial f(z;\cdot)}{\partial \overline{z}}\right\}$ are straightforward to obtain given the formula for $f(z;0,\underline{z},\hat{z},\overline{z})$:

$$f(z;0,\underline{z},\hat{z},\overline{z}) = \begin{cases} 2\frac{z-\underline{z}}{(\hat{z}-\underline{z})(\overline{z}-\underline{z})} & \text{for } z < \hat{z} \\ 2\frac{\overline{z}-z}{(\overline{z}-\hat{z})(\overline{z}-\underline{z})} & \text{for } z \geq \hat{z} \end{cases}$$

Differentiating $f(z;0,\underline{z},\hat{z},\overline{z})$ with respect to $\underline{z}$, $\hat{z}$ and $\overline{z}$, and evaluating at the optimal policy $\{\underline{z},\hat{z},\overline{z}\} = \{-\overline{z}_0, 0, \overline{z}_0\}$ yields for $z < 0$:

$$\left\{\frac{\partial f(z;\cdot)}{\partial \underline{z}}, \frac{\partial f(z;\cdot)}{\partial \hat{z}}, \frac{\partial f(z;\cdot)}{\partial \overline{z}}\right\}\Bigg|_{\mu=0,\underline{z}=-\overline{z}_0,\hat{z}=0,\overline{z}=\overline{z}_0} = \left\{\frac{3z+\overline{z}_0}{2\overline{z}_0^3}, -\frac{z+\overline{z}}{\overline{z}_0^3}, -\frac{z+\overline{z}}{2\overline{z}_0^3}\right\}$$

The derivative with respect to μ is somewhat more complicated. First, set policy variables to their optimal values under $\mu = 0$:

$$f(z;\mu,-\overline{z}_0,0,\overline{z}_0) = \frac{1 - v(\overline{z}_0,\mu) + v(-z,\mu) - v(-\overline{z}_0 - z,\mu)}{2\overline{z}_0 - \overline{z}_0 v(-\overline{z}_0,\mu) - \overline{z}_0 v(\overline{z}_0,\mu)} \text{ for } z < 0$$

Second, denote the numerator by $N(\mu)$ and denominator by $D(\mu)$, so that:

$$\frac{\partial f(z;\mu,-\overline{z}_0,0,\overline{z}_0)}{\partial \mu} = \frac{N'(\mu)D(\mu) - D'(\mu)N(\mu)}{D(\mu)^2} \text{ for } z < 0 \tag{A.5}$$

Third, note that $v_\mu(z,\mu) = \frac{2}{\sigma^2} z v(z,\mu)$ and thus derivatives of numerator and denominator are given by:

$$N^k(\mu) = \frac{2^k}{\sigma^{2k}}\Big(- \overline{z}_0^k v(\overline{z}_0,\mu) + (-z)^k v(-z,\mu) - (-\overline{z}_0 - z)^k v(-\overline{z}_0 - z,\mu)\Big)$$
$$D^k(\mu) = \frac{2^k}{\sigma^{2k}}\Big((-\overline{z}_0)^{k+1} v(-\overline{z}_0,\mu) - \overline{z}_0^{k+1} v(\overline{z}_0,\mu)\Big)$$

Since $v(z,0) = 1$, evaluating at $\mu = 0$ yields:

$$N^k(0) = \begin{cases} \frac{2^k}{\sigma^{2k}}\Big(- \overline{z}_0^k - z^k + (z+\overline{z}_0)^k\Big) & \text{for } k \text{ odd} \\ \frac{2^k}{\sigma^{2k}}\Big(- \overline{z}_0^k + z^k - (z+\overline{z}_0)^k\Big) & \text{for } k \text{ even} \end{cases}$$

$$D^k(0) = \begin{cases} 0 & \text{for } k \text{ odd} \\ \frac{2^k}{\sigma^{2k}}\Big(- 2\overline{z}_0^{k+1}\Big) & \text{for } k \text{ even} \end{cases}$$

Note that $N(0) = D(0) = 0$ and it follows that evaluating (A.5) at $\mu = 0$ directly is not possible since both the numerator and the denominator converge to zero as $\mu \to 0$. Applying L'Hospital's rule four times yields:

$$\left.\frac{\partial f(z;\mu,\overline{z}_0,0,\overline{z}_0)}{\partial\mu}\right|_{\mu=0} = -\frac{z^2+\overline{z}_0 z}{\sigma^2\overline{z}_0^2} \qquad \text{for } z<0$$

Finally, collecting all the terms:

$$\left.\frac{\partial f(z,\mu)}{\partial\mu}\right|_{\mu=0} = -\frac{z^2+\overline{z}_0 z}{\sigma^2\overline{z}_0^2} + \frac{z}{\overline{z}_0^3}\frac{\partial\overline{z}(0)}{\partial\mu} - \frac{z+\overline{z}_0}{\overline{z}_0^3}\frac{\partial\hat{z}(0)}{\partial\mu} \quad \text{for } z<0$$

which provides derivative of density with respect to drift μ at any point $z \in [-\overline{z}_0, 0)$. Density is non-differentiable at $z = 0$ and for positive values $z \in (0, \overline{z}_0]$ derivative of density is given by $\frac{\partial f(z,0)}{\partial\mu} = -\frac{\partial f(-z,\mu)}{\partial\mu}$.

A.1.5 Impact effect under non-zero drift

Recall that for a positive shock $\delta > 0$, impact effect is given by:

$$\Theta(\delta,\mu) = \int_{\underline{z}(\mu)-\delta}^{\underline{z}(\mu)} (\hat{z}(\mu)-z) f(z+\delta,\mu)dz$$

and its derivative with respect to μ is:

$$\begin{aligned}\frac{\partial\Theta(\delta,\mu)}{\partial\mu} = &\frac{\partial\underline{z}(\mu)}{\partial\mu}\Delta^+(\mu)f(\underline{z}(\mu)+\delta,\mu)+ \\ &\int_{\underline{z}(\mu)-\delta}^{\underline{z}(\mu)}\left(\frac{\partial\hat{z}(\mu)}{\partial\mu}f(z+\delta,\mu) + (\hat{z}(\mu)-z)\frac{\partial f(z+\delta,\mu)}{\partial\mu}\right)dz\end{aligned}$$

where $\Delta^+(\mu) = \hat{z}(\mu) - \underline{z}(\mu)$ and I have used the fact that $f(\underline{z}(\mu),\mu) = 0$. Evaluating at $\mu = 0$ yields:

$$\begin{aligned}\frac{\partial\Theta(\delta,0)}{\partial\mu} &= \overline{z}_0\frac{\partial\underline{z}(0)}{\partial\mu}f(-\overline{z}_0+\delta,0) + \int_{-\overline{z}_0-\delta}^{-\overline{z}_0}\left(\frac{\partial\hat{z}(\mu)}{\partial\mu}f(z+\delta,0) - z\frac{\partial f(z+\delta,0)}{\partial\mu}\right)dz \\ &= \overline{z}_0\frac{\partial\underline{z}(0)}{\partial\mu}f(-\overline{z}_0+\delta,0) + \frac{\partial\hat{z}(\mu)}{\partial\mu}F(-\overline{z}_0+\delta) - \int_{-\overline{z}_0-\delta}^{-\overline{z}_0} z\frac{\partial f(z+\delta,0)}{\partial\mu}dz\end{aligned}$$

Previous sections of Appendix provide expressions for all terms in the above equation. Note that as long as $\delta < \overline{z}_0$, the integral in the last term is well defined, however if $\delta \geq \overline{z}_0$, then it has to be split into two integrals since $f(z,0)$ is not differentiable at $z = 0$:

$$\int_{-\overline{z}_0-\delta}^{-\overline{z}_0} z\frac{\partial f(z+\delta,0)}{\partial\mu}dz = \int_{-\overline{z}_0-\delta}^{-\delta} z\frac{\partial f(z+\delta,0)}{\partial\mu}dz + \int_{-\delta}^{-\overline{z}_0} z\frac{\partial f(z+\delta,0)}{\partial\mu}dz$$

A direct computation provides the following result:

$$\left.\frac{\partial\Theta(\delta,\mu)}{\partial\mu}\right|_{\mu=0} = \begin{cases} \frac{\delta^2(6\overline{z}_0^2-\delta^2-2\delta\overline{z}_0)}{12\sigma^2\overline{z}_0^2} - \frac{\delta^3}{6\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu}, & \text{for } \delta < \overline{z}_0 \\ \frac{\delta^2(\delta^2-2\delta\overline{z}_0-6\overline{z}_0^2)+\overline{z}_0^3(16\delta-6\overline{z}_0)}{12\sigma^2\overline{z}_0^2} - \frac{\delta^3-12\delta\overline{z}_0^2+12\overline{z}_0^3}{6\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu}, & \text{for } \delta \in [\overline{z}_0, 2\overline{z}_0) \\ \frac{\overline{z}_0^2}{6\sigma^2} + \frac{2}{3}\frac{\partial\Delta^+(0)}{\partial\mu}, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}$$

A.1.6 Cumulative Impulse Response under non-zero drift

It is convenient to split $M(\delta,\mu)$ into several smaller parts. First, note that function $m(z,\mu)$ can be written as:

$$\begin{aligned} m(z,\mu) &= -\mathbb{E}\left(\int_0^\tau (z(s)-\bar{x}(\mu)ds \;\middle|\; z(0)=z\right) \\ &= \underbrace{-\mathbb{E}\left(\int_0^\tau z(s)ds \;\middle|\; z(0)=z\right)}_{\hat{m}(z,\mu)} + \bar{x}(\mu)\underbrace{\mathbb{E}\left(\tau \;\middle|\; z(0)=z\right)}_{\tau(z,\mu)} \\ &= \hat{m}(z,\mu) + \bar{x}(\mu)\tau(z,\mu) \end{aligned}$$

Function $\hat{m}(z,\mu)$ is now the expected cumulative gap until first adjustment and is defined by the following ODE:

$$z = -\mu\hat{m}_z(z,\mu) + \frac{\sigma^2}{2}\hat{m}_{zz}(z,\mu) \tag{A.6}$$

with boundary conditions $\hat{m}(\underline{z}(\mu),\mu) = \hat{m}(\overline{z}(\mu),\mu) = 0$. Function $\tau(z,\mu)$ is the expected time of first adjustment conditional on $z(0) = z$, and is also defined by ODE:

$$0 = 1 - \mu\tau_z(z,\mu) + \frac{\sigma^2}{2}\tau_{zz}(z,\mu)$$

and boundary conditions $\tau(\underline{z}(\mu),\mu) = \tau(\overline{z}(\mu),\mu) = 0$. Solution to (A.6) is:

$$\hat{m}(z,\mu) = C_1 + C_2 e^{\frac{2\mu}{\sigma^2}z} - \frac{1}{2\mu}z^2 - \frac{\sigma^2}{2\mu^2}z \tag{A.7}$$

where C_1 and C_2 are determined by boundary conditions. Solution to $\tau(z,\mu)$ is provided in Chapter 5.5 of Stokey (2009).

Using this notation, express $M(\delta,\mu)$ as follows:

$$\begin{aligned}
M(\delta,\mu) &= \int_{\underline{z}(\mu)}^{\overline{z}(\mu)-\delta} m(z,\mu) f(z+\delta,\mu) dz \;-\; \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} m(z,\mu) f(z,\mu) dz \\
&= \underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(\mu)-\delta} \hat{m}(z,\mu) f(z+\delta,\mu) dz}_{\hat{M}(\delta,\mu)} \;+\; \bar{x}(\mu) \underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(\mu)-\delta} \tau(z,\mu) f(z+\delta,\mu) dz}_{T(\delta,\mu)} \\
&\quad - \underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \hat{m}(z,\mu) f(z,\mu) dz}_{\hat{M}(0,\mu)} \;-\; \bar{x}(\mu) \underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \tau(z,\mu) f(z,\mu) dz}_{T(0,\mu)} \\
&= \hat{M}(\delta,\mu) - \hat{M}(0,\mu) \;+\; \bar{x}(\mu)\big[T(\delta,\mu) - T(0,\mu)\big]
\end{aligned}$$

and thus:

$$\begin{aligned}
\frac{\partial M(\delta,0)}{\partial\mu} &= \frac{\partial \hat{M}(\delta,0)}{\partial\mu} - \frac{\partial \hat{M}(0,0)}{\partial\mu} + \frac{\partial \bar{x}(0)}{\partial\mu}\big[T(\delta,0) - T(0,0)\big] \\
&\quad + \underbrace{\bar{x}(0)}_{=0}\Big[\frac{\partial T(\delta,0)}{\partial\mu} - \frac{\partial T(0,0)}{\partial\mu}\Big]
\end{aligned}$$

where $\bar{x}(0) = 0$ due to symmetry of $f(z,0)$. Derivatives of $\hat{M}(\delta,\mu)$ and $\bar{x}(\mu)$ are given by:

$$\begin{aligned}
\frac{\partial \hat{M}(\delta,0)}{\partial\mu} &= \frac{\partial \overline{z}(0)}{\partial\mu} \hat{m}(\overline{z}(0)-\delta,0) \overbrace{f(\overline{z}(0),0)}^{=0} \;-\; \frac{\partial \underline{z}(0)}{\partial\mu} \overbrace{\hat{m}(\underline{z}(0),0)}^{=0} f(\underline{z}(0)+\delta,0) \\
&\quad + \int_{\underline{z}(0)}^{\overline{z}(0)-\delta} \frac{\partial \hat{m}(z,0)}{\partial\mu} f(z+\delta,0) dz \;+\; \int_{\underline{z}(0)}^{\overline{z}(0)-\delta} \hat{m}(z,0) \frac{\partial f(z+\delta,0)}{\partial\mu} dz
\end{aligned}$$

$$\frac{\partial \bar{x}(0)}{\partial\mu} = \frac{\partial \overline{z}(0)}{\partial\mu} \overline{z}(0) \underbrace{f(\overline{z}(0),0)}_{=0} \;-\; \frac{\partial \underline{z}(0)}{\partial\mu} \underline{z}(0) \underbrace{f(\underline{z}(0),0)}_{=0} \;+\; \int_{\underline{z}(0)}^{\overline{z}(0)} z \frac{\partial f(z,0)}{\partial\mu} dz$$

where derivatives of integration boundaries are zero due to boundary conditions of $\hat{m}(z,0)$ and $f(z,0)$. Note that integrals have to split accordingly since stationary density $f(z,0)$ is not differentiable at $z = 0$. Derivative of stationary density $f(z,\mu)$ with respect to drift is provided in Appendix A.1.4. A direct computation yields:

$$\frac{\partial \bar{x}(0)}{\partial\mu} = \frac{2}{3}\frac{\partial \overline{z}(0)}{\partial\mu} + \frac{1}{3}\frac{\partial \hat{z}(0)}{\partial\mu} - \frac{\overline{z}_0^2}{6\sigma^2}$$

Stokey (2009) shows that $\tau(z,0)=\frac{\overline{z}_0^2-z^2}{\sigma^2}$ and thus computing $T(\delta,0)$ gives:

$$T(\delta,0)=\begin{cases}\frac{1}{12\sigma^2\overline{z}_0^2}\left[\delta^4+4\delta^3\overline{z}_0-12\delta^2\overline{z}_0^2+10\overline{z}_0^4\right], & \text{for } \delta<\overline{z}_0\\ \frac{1}{12\sigma^2\overline{z}_0^2}\left[-\delta^4+4\delta^3\overline{z}_0-16\delta\overline{z}_0^3+16\overline{z}_0^4\right], & \text{for } \delta\in[\overline{z}_0,2\overline{z}_0)\\ 0, & \text{for } \delta\geq 2\overline{z}_0\end{cases}$$

Computation of $\frac{\partial\hat{M}(\delta,0)}{\partial\mu}$ requires knowledge of $\frac{\partial\hat{m}(z,0)}{\partial\mu}$ and $\hat{m}(z,0)$. The latter solves (A.6) for $\mu=0$ and is given by:

$$\hat{m}(z,0)=\frac{z^3-\overline{z}_0^2z}{3\sigma^2}$$

It remains to characterize derivative of $\hat{m}(z,\mu)$ with respect to μ and then derivative of CIR can be computed. First, note that $\hat{m}(z,\mu)$ depends on μ both directly as can be seen in (A.7), as well as indirectly through boundaries of inaction region that appear in expressions for C_1 and C_2:

$$C_2=\frac{1}{v(\overline{z}(\mu),\mu)-v(\underline{z}(\mu),\mu)}\left[\frac{1}{2\mu}\left(\overline{z}(\mu)^2-\underline{z}(\mu)^2\right)+\frac{\sigma^2}{2\mu^2}\left(\overline{z}(\mu)-\underline{z}(\mu)\right)\right]$$

$$C_1=\frac{1}{2\mu}\overline{z}(\mu)^2+\frac{\sigma^2}{2\mu^2}\overline{z}(\mu)-C_2v(\overline{z}(\mu),\mu)$$

where $v(z,\mu)=e^{\frac{2\mu}{\sigma^2}z}$. It is thus convenient to include the boundaries explicitly as arguments of $\hat{m}(z,\mu)$, so that $\hat{m}(z,\mu)=\hat{m}(z;\mu,\overline{z}(\mu),\underline{z}(\mu))$. Then:

$$\begin{aligned}\left.\frac{\partial\hat{m}(z,\mu)}{\partial\mu}\right|_{\mu=0}&=\left.\frac{d\hat{m}(z;\mu,\overline{z},\underline{z})}{d\mu}\right|_{\mu=0,\overline{z}=\overline{z}_0,\underline{z}=-\overline{z}_0}\\&=\left.\left(\frac{\partial\hat{m}(z;\cdot)}{\partial\mu}+\frac{\partial\hat{m}(z;\cdot)}{\partial\overline{z}}\frac{\partial\overline{z}(\mu)}{\partial\mu}+\frac{\partial\hat{m}(z;\cdot)}{\partial\underline{z}}\frac{\partial\underline{z}(\mu)}{\partial\mu}\right)\right|_{\mu=0,\overline{z}=\overline{z}_0,\underline{z}=-\overline{z}_0}\end{aligned}$$

Derivatives with respect to boundaries $\underline{z}$ and $\overline{z}$ are relatively easy to obtain. Set $\mu=0$, then:

$$\hat{m}(z;0,\overline{z},\underline{z})=\frac{z^3-\underline{z}^3}{3\sigma^2}-\frac{\overline{z}^3-\underline{z}^3}{3\sigma^2(\overline{z}-\underline{z})}(z-\underline{z})$$

$$\left.\frac{\partial\hat{m}(z;\mu,\overline{z},\underline{z})}{\partial\overline{z}}\right|_{\mu=0,\overline{z}=\overline{z}_0,\underline{z}=-\overline{z}_0}=-\frac{\overline{z}_0z+\overline{z}_0^2}{3\sigma^2}$$

$$\left.\frac{\partial\hat{m}(z;\mu,\overline{z},\underline{z})}{\partial\underline{z}}\right|_{\mu=0,\overline{z}=\overline{z}_0,\underline{z}=-\overline{z}_0}=\frac{\overline{z}_0z-\overline{z}_0^2}{3\sigma^2}$$

Obtaining derivative of $\hat{m}(z, \mu)$ is somewhat more involved. Setting $\overline{z} = \overline{z}_0, \underline{z} = -\overline{z}_0$:

$$\hat{m}(z; \mu, \overline{z}_0, -\overline{z}_0) = \frac{2\overline{z}_0\sigma^2\beta(z,\mu) + \mu\gamma(\mu)(\overline{z}_0^2 - z^2) + \sigma^2\gamma(\mu)(\overline{z}_0 - z)}{2\mu^2\gamma(\mu)}$$

where $\gamma(\mu) = v(\overline{z}_0, \mu) - v(-\overline{z}_0, \mu)$ and $\beta(z, \mu) = v(z, \mu) - v(\overline{z}_0, \mu)$. Differentiating with respect to μ and collecting terms:

$$\left.\frac{\partial \hat{m}(z; \mu, \overline{z}, \underline{z})}{\partial \mu}\right|_{\mu=0,\overline{z}=\overline{z}_0,\underline{z}=-\overline{z}_0} = \tag{A.8}$$

$$= \frac{2\overline{z}_0\sigma^2\big(\mu\gamma(\mu)\beta'_\mu(z,\mu) - 2\gamma(\mu)\beta(z,\mu) - \mu\gamma'(\mu)\beta(z,\mu)\big) - \mu\gamma(\mu)^2(\overline{z}_0^2 - z^2) - 2\sigma^2\gamma(\mu)^2(\overline{z}_0 - z)}{2\mu^3\gamma(\mu)^2}$$

Note that as $\mu \to 0$, $\gamma(\mu) \to 0$ and $\beta(z, \mu) \to 0$. In addition, derivatives of $\gamma(\mu)$ and $\beta(z, \mu)$ with respect to μ evaluated at $\mu = 0$ are given by:

$$\gamma^k(0) = \begin{cases} \frac{2^{k+1}}{\sigma^{2k}}\overline{z}_0^k & \text{for } k \text{ odd} \\ 0 & \text{for } k \text{ even} \end{cases}$$

$$\beta^k_\mu(z, 0) = \frac{2^k}{\sigma^{2k}}(z^k - \overline{z}_0^k)$$

This implies that evaluating (A.8) at $\mu = 0$ is not possible as both denominator and numerator are zero at $\mu = 0$. Applying L'Hospital's rule five times provides the result:

$$\left.\frac{\partial \hat{m}(z; \mu, \overline{z}, \underline{z})}{\partial \mu}\right|_{\mu=0,\overline{z}=\overline{z}_0,\underline{z}=-\overline{z}_0} = \frac{(\overline{z}_0^2 - z^2)^2}{6\sigma^4}$$

Collecting all the terms gives the derivative of interest:

$$\left.\frac{\partial \hat{m}(z, \mu)}{\partial \mu}\right|_{\mu=0} = \frac{(\overline{z}_0^2 - z^2)^2}{6\sigma^4} - \frac{2\overline{z}_0^2}{3\sigma^2}\frac{\partial \underline{z}(\mu)}{\partial \mu}$$

Now all necessary ingredients for the derivative of cumulative impulse response with respect to drift μ are collected and direct computation yields:

$$\left.\frac{\partial M(\delta, \mu)}{\partial \mu}\right|_{\mu=0} = \begin{cases} \frac{1}{360\sigma^4\overline{z}_0^2}\left[-4\delta^6 - 18\delta^5\overline{z}_0 + 45\delta^4\overline{z}_0^2 + 20\delta^3\overline{z}_0^3 - 60\delta^2\overline{z}_0^4\right] \\ -\frac{1}{180\sigma^2\overline{z}_0^3}\left[3\delta^5 + 10\delta^4\overline{z}_0\right]\frac{\partial\Delta^+(0)}{\partial\mu}, & \text{for } \delta < \overline{z}_0 \\ \frac{1}{360\sigma^4\overline{z}_0^2}\left[4\delta^6 - 18\delta^5\overline{z}_0 + 15\delta^4\overline{z}_0^2 + 20\delta^3\overline{z}_0^3 - 48\delta\overline{z}_0^5 + 10\overline{z}_0^6\right] \\ -\frac{1}{180\sigma^2\overline{z}_0^3}\left[3\delta^5 - 10\delta^4\overline{z}_0 + 80\delta\overline{z}_0^4 - 60\overline{z}_0^5\right]\frac{\partial\Delta^+(0)}{\partial\mu}, & \text{for } \delta \in [\overline{z}_0, 2\overline{z}_0) \\ -\frac{\overline{z}_0^4}{60\sigma^4} - \frac{\overline{z}_0^2}{5\sigma^2}\frac{\partial\Delta^+(0)}{\partial\mu}, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}$$

A.1.7 CIR for a Shock to the Drift

As in Appendix A.1.6, rewrite $m(z,\mu)$ as follows:

$$\begin{aligned}
m(z,\mu) &= \mathbb{E}\left(\int_0^\tau (z(s)-\bar{x}(\mu)ds \,\middle|\, z(0)=z\right)\\
&= \underbrace{\mathbb{E}\left(\int_0^\tau z(s)ds \,\middle|\, z(0)=z\right)}_{\hat{m}(z,\mu)} - \bar{x}(\mu)\underbrace{\mathbb{E}\left(\tau \,\middle|\, z(0)=z\right)}_{\tau(z,\mu)}\\
&= \hat{m}(z,\mu) - \bar{x}(\mu)\tau(z,\mu)
\end{aligned}$$

with $\hat{m}(z,\mu)$ being defined by the following ODE:

$$0 = z - \mu\hat{m}_z(z,\mu) + \frac{\sigma^2}{2}\hat{m}_{zz}(z,\mu)$$

with boundary conditions $\hat{m}(\underline{z}(\mu),\mu) = \hat{m}(\overline{z}(\mu),\mu) = 0$. This is the exact same function as in Appendix A.1.6, multiplied by -1. Similarly, and with some abuse of notation:

$$\begin{aligned}
M(\mu) &= \int_{\underline{z}(\mu)}^{\overline{z}(0)} m(z,\mu)f(z,0)dz \;-\; \int_{\underline{z}(\mu)}^{\overline{z}(\mu)} m(z,\mu)f(z,\mu)dz\\
&= \underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(0)} \hat{m}(z,\mu)f(z,0)dz}_{\hat{M}_0(\mu)} \;-\; \bar{x}(\mu)\underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(0)} \tau(z,\mu)f(z,0)dz}_{T_0(\mu)}\\
&\quad - \underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \hat{m}(z,\mu)f(z,\mu)dz}_{\hat{M}(\mu)} \;+\; \bar{x}(\mu)\underbrace{\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} \tau(z,\mu)f(z,\mu)dz}_{T(\mu)}\\
&= \hat{M}_0(\mu) - \hat{M}(\mu) \;-\; \bar{x}(\mu)\big[T_0(\mu) - T(\mu)\big]
\end{aligned}$$

and thus:

$$\begin{aligned}
\frac{\partial M(0)}{\partial\mu} &= \frac{\partial\hat{M}_0(0)}{\partial\mu} - \frac{\partial\hat{M}(0)}{\partial\mu} - \frac{\partial\bar{x}(0)}{\partial\mu}\underbrace{\big[T_0(0)-T(0)\big]}_{=0}\\
&\quad - \underbrace{\bar{x}(0)}_{=0}\left[\frac{\partial T_0(0)}{\partial\mu} - \frac{\partial T(0)}{\partial\mu}\right]
\end{aligned}$$

Derivatives of $\hat{M}_0(\mu)$ and $\hat{M}(\mu)$ are given by:

$$\frac{\partial \hat{M}_0(0)}{\partial \mu} = -\frac{\partial \underline{z}(0)}{\partial \mu}\overbrace{\hat{m}(\underline{z}(0),0)f(\underline{z}(0),0)}^{=0} + \int_{\underline{z}(0)}^{\overline{z}(0)} \frac{\partial \hat{m}(z,0)}{\partial \mu} f(z,0)dz$$

$$\frac{\partial \hat{M}(0)}{\partial \mu} = \frac{\partial \overline{z}(0)}{\partial \mu}\overbrace{\hat{m}(\overline{z}(0),0)f(\overline{z}(0),0)}^{=0} - \frac{\partial \underline{z}(0)}{\partial \mu}\overbrace{\hat{m}(\underline{z}(0),0)f(\underline{z}(0),0)}^{=0}$$
$$+ \int_{\underline{z}(0)}^{\overline{z}(0)} \frac{\partial \hat{m}(z,0)}{\partial \mu} f(z,0)dz + \int_{\underline{z}(0)}^{\overline{z}(0)} \hat{m}(z,0)\frac{\partial f(z,0)}{\partial \mu}dz$$

so that:

$$\frac{\partial M(0)}{\partial \mu} = -\int_{\underline{z}(0)}^{\overline{z}(0)} \hat{m}(z,0)\frac{\partial f(z,0)}{\partial \mu}dz$$

Using results from Appendix A.1.6, one can obtains that:

$$\frac{\partial M(0)}{\partial \mu} = \frac{\overline{z}_0^2}{180\sigma^2}\left[\frac{6\overline{z}_0^2}{\sigma^2} - 16\frac{\partial \underline{z}(0)}{\partial \mu} - 14\frac{\partial \hat{z}(0)}{\partial \mu}\right]$$

Lemma 3 in Appendix A.2.10 shows that this expression is positive.

A.1.8 Random Opportunities of Costless Adjustment

This section illustrates how the results of this paper can be extended to a model that allows for random costless adjustments. This class of models is usually referred to as "CalvoPlus" in the literature, as it nests both the traditional Calvo (1983) model and the standard menu cost model. Such an extension provides a more realistic distribution of price adjustments and allows for a better fit to the data.

Assume that the problem of a firm is exactly as described in section 1.2.1, with the only difference that the firm occasionally gets an opportunity to adjust its price at no cost. These opportunities arrive at Poisson rate $\lambda > 0$. The value function of the firm now satisfies the following HJB equation:

$$\rho v(z) = -z^2 - \mu v'(z) + \frac{\sigma^2}{2}v''(z) + \lambda(v(\hat{z}) - v(z))$$

together with the same set of optimality and smoothness conditions: $v'(\underline{z}) = v'(\overline{z}) = v'(\hat{z}) = 0$, $v(\underline{z}) = v(\overline{z}) = v(\hat{z}) - \kappa$. Rewrite the HJB equation in the following way:

$$\underbrace{(\rho + \lambda)}_{\hat{\rho}} v(z) = -z^2 - \mu v'(z) + \frac{\sigma^2}{2}v''(z) + \lambda v(\hat{z})$$

and note that since $\lambda v(\hat{z})$ is a constant, it drops out in all optimality and smoothness conditions, so that the results of Proposition 1 regarding the effect of drift on firms' optimal policy apply. In addition, the expressions for the derivatives of the reset point and inaction region boundaries, provided in Appendix A.1.3, are unchanged as well, with the only difference that ρ is substituted with $\hat{\rho} = \rho + \lambda$.

Stationary density is now determined by the following KFE:

$$0 = -\lambda f(z,\mu) + \mu f_z(z,\mu) + \frac{\sigma^2}{2} f_{zz}(z,\mu)$$

together with boundary conditions $f(\underline{z}(\mu),\mu) = f(\overline{z}(\mu),\mu) = 0$, unit mass condition $\int_{\underline{z}(\mu)}^{\overline{z}(\mu)} f(z,\mu)dz = 1$ and continuity at $z = \hat{z}(\mu)$. Stationary density takes the following form:

$$f(z,\mu) = \begin{cases} C_1 e^{R_1 z} + C_2 e^{R_2 z} & \text{for } z < \hat{z} \\ C_3 e^{R_1 z} + C_4 e^{R_2 z} & \text{for } z \geq \hat{z} \end{cases}$$

where $R_1 = \frac{-\mu - \sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2}$, $R_2 = \frac{-\mu + \sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2}$ and coefficients C_1, C_2, C_3, C_4 are determined by the four above listed conditions on $f(z,\mu)$.

Similarly, the expressions for the expected (negative) cumulative gap until first adjustment $\hat{m}(z,\mu) = -\mathbb{E}\left(\int_0^\tau z(s)ds \;\middle|\; z(0) = z\right)$ and the expected time of adjustment $\tau(z,\mu) = \mathbb{E}\left(\tau \;\middle|\; z(0) = z\right)$ have to be adjusted. They are now defined by the following ODEs:

$$z = -\lambda m(z,\mu) - \mu \hat{m}_z(z,\mu) + \frac{\sigma^2}{2}\hat{m}_{zz}(z,\mu)$$
$$0 = 1 - \lambda\tau(z,\mu) - \mu\tau_z(z,\mu) + \frac{\sigma^2}{2}\tau_{zz}(z,\mu)$$

with boundary conditions $\hat{m}(\underline{z}(\mu),\mu) = \hat{m}(\overline{z}(\mu),\mu) = 0$ and $\tau(\underline{z}(\mu),\mu) = \tau(\overline{z}(\mu),\mu) = 0$.

Obtaining derivatives of $\Theta(\delta,\mu)$ and $M(\delta,\mu)$ with respect to drift requires derivatives of $f(z,\mu)$ and $m(\hat{z},\mu)$, as well as expressions for $f(z,\mu)$, $m(\hat{z},\mu)$ and $\tau(z,\mu)$ evaluated at $\mu = 0$. The former can be obtained by using the Implicit Function Theorem, as in Appendix A.1.3, and are given by:

$$\left.\frac{\partial f(z,\mu)}{\partial \mu}\right|_{\mu=0} = \frac{1}{2(q(\overline{z}_0)-2)}\left[-\frac{\alpha z p(\overline{z}_0+z)}{\sigma^2} + \frac{2\alpha^2 p(z)}{p(\overline{z}_0)}\frac{\partial \underline{z}(0)}{\partial \mu} - \frac{\alpha^2 q(\overline{z}_0)p(\overline{z}_0+z)}{p(\overline{z}_0)}\frac{\partial \hat{z}(0)}{\partial \mu}\right], \text{ if } z \leq 0$$

$$\left.\frac{\partial \hat{m}(z,\mu)}{\partial \mu}\right|_{\mu=0} = \frac{\overline{z}_0 z p(z)}{\lambda\sigma^2 p(\overline{z}_0)} + \frac{1}{\lambda^2} - \frac{q(z)}{q(\overline{z}_0)}\left[\frac{\alpha^2\overline{z}_0^2 + 2}{2\lambda^2} + \frac{\alpha\overline{z}_0 q(\overline{z}_0) - p(\overline{z}_0)}{\lambda p(\overline{z}_0)}\frac{\partial \underline{z}(0)}{\partial \mu}\right]$$

where $\alpha = \sqrt{2\lambda}/\sigma$, $p(z) = e^{\alpha z} - e^{-\alpha z}$, $q(z) = e^{\alpha z} + e^{-\alpha z}$, and the expression for $\frac{\partial f(z,\mu)}{\partial \mu}\big|_{\mu=0}$ for $z > 0$ can be obtained by using the fact that $\frac{\partial f(z,\mu)}{\partial \mu}\big|_{\mu=0} = -\frac{\partial f(-z,\mu)}{\partial \mu}\big|_{\mu=0}$.

The expressions for $f(z,0)$, $m(\hat{z},0)$ and $\tau(z,0)$ are given by:

$$
\begin{aligned}
f(z,0) &= \frac{\alpha}{2(q(\overline{z}_0)-2)} p(\overline{z}_0 - |z|) \\
\hat{m}(z,0) &= \frac{\overline{z}_0 p(z) - z p(\overline{z}_0)}{\lambda p(\overline{z}_0)} \\
\tau(z,0) &= \frac{q(\overline{z}_0) - q(z)}{\lambda q(\overline{z}_0)}
\end{aligned}
$$

Whereas it is quite challenging to obtain analogues of Propositions 2 and 5 regarding the effects of μ on impact and cumulative responses for shocks of arbitrary size, one can still show that the overshooting result of Proposition 6 extends to the case with random costless adjustments. It suffices to show that $\frac{\partial \Theta(2\overline{z}_0,\mu)}{\partial \mu}\big|_{\mu=0} > 0$ and $\frac{\partial M(2\overline{z}_0,\mu)}{\partial \mu}\big|_{\mu=0} < 0$. The expressions for these derivatives can be obtained using the expressions provided above and are given by:

$$
\frac{\partial \Theta(2\overline{z}_0,\mu)}{\partial \mu}\bigg|_{\mu=0} = \frac{1}{2(q(\overline{z}_0)-2)}\left[\frac{2q(\overline{z}_0)-4-2\alpha^2\overline{z}_0^2}{\lambda} + \frac{4(\alpha\overline{z}_0 q(\overline{z}_0)-p(\overline{z}_0))}{p(\overline{z}_0)}\left(\frac{\partial \hat{z}(0)}{\partial \mu} - \frac{\partial \underline{z}(0)}{\partial \mu}\right)\right]
$$

$$
\frac{\partial M(2\overline{z}_0,\mu)}{\partial \mu}\bigg|_{\mu=0} = -\frac{2q(2\overline{z}_0) - 8q(\overline{z}_0) + 12 - \alpha^3\overline{z}_0^3 p(\overline{z}_0)}{2\lambda^2(q(\overline{z}_0)-2)^2}
$$

$$
-\frac{\alpha^2\overline{z}_0^2 q(3\overline{z}_0) - 2\alpha\overline{z}_0 p(3\overline{z}_0) + 2\alpha\overline{z}_0 p(2\overline{z}_0) + 3\alpha^2\overline{z}_0^2 q(\overline{z}_0) + 2\alpha\overline{z}_0 p(\overline{z}_0) - 8\alpha^2\overline{z}_0^2}{2\lambda p(\overline{z}_0)^2(q(\overline{z}_0)-2)^2}\left(\frac{\partial \hat{z}(0)}{\partial \mu} - \frac{\partial \underline{z}(0)}{\partial \mu}\right)
$$

and the fact that $\frac{\partial \Theta(2\overline{z}_0,\mu)}{\partial \mu}\big|_{\mu=0} > 0$ and $\frac{\partial M(2\overline{z}_0,\mu)}{\partial \mu}\big|_{\mu=0} < 0$ is proven in Lemma 4 in Appendix A.2.11.

A.2 Proofs

A.2.1 Some useful expressions

I provide several expressions which will be used later. Let $w_1 = e^x - e^{-x}$ and $w_2 = e^x + e^{-x}$ where $x > 0$. Using Tailor expansion one obtains following results:

$$w_1 = 2 \sum_{i=1,3,5...}^{\infty} \frac{x^i}{i!} > 0$$

$$w_2 = 2 + 2 \sum_{i=2,4,6...}^{\infty} \frac{x^i}{i!} > w_1$$

$$w_1 w_2 = e^{2x} - e^{-2x} = 2 \sum_{i=1,3,5...}^{\infty} \frac{2^i x^i}{i!}$$

$$w_1^2 = e^{2x} + e^{-2x} - 2 = 2 \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{i!}$$

$$w_2^2 = e^{2x} + e^{-2x} + 2 = 4 + 2 \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{i!}$$

$$w_1^3 = e^{3x} - e^{-3x} - 3w_1 = 2 \sum_{i=1,3,5...}^{\infty} \frac{3^i x^i}{i!} - 3w_1$$

$$w_1^2 w_2 = e^{3x} + e^{-3x} - w_2 = 2 + 2 \sum_{i=2,4,6...}^{\infty} \frac{3^i x^i}{i!} - w_2$$

A.2.2 Proof of Proposition 1

First, let's show that $\frac{\partial \overline{z}}{\partial \mu}\big|_{\mu=0} > 0$. Denote $x := \alpha \overline{z}_0$:

$$\left.\frac{\partial \overline{z}}{\partial \mu}\right|_{\mu=0} = \frac{4x^2 + x w_1 w_2 - 2w_1^2}{2\rho(x w_1 w_2 - w_1^2)}$$

Firstly, using expressions from Appendix A.2.1, one can show that denominator is positive:

$$2\rho(x w_1 w_2 - w_1^2) > 0 \iff x w_2 - w_1 > 0 \iff \sum_{i=3,5,7...}^{\infty} \frac{x^i}{(i-1)!} - \sum_{i=3,5,7...}^{\infty} \frac{x^i}{i!} > 0$$

where last inequality is trivially satisfied. Secondly, similar logic applies to the numerator:

$$4x^2 + x w_1 w_2 - 2w_1^2 > 0 \iff 4x^2 + \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{(i-1)!} - \sum_{i=2,4,6...}^{\infty} \frac{2^{i+2} x^i}{i!} > 0$$

$$\iff \sum_{i=6,8,10...}^{\infty} \frac{2^i x^i}{(i-1)!} - \sum_{i=6,8,10...}^{\infty} \frac{2^{i+2} x^i}{i!} > 0$$

where last line follows since $\frac{2^i}{(i-1)!} > \frac{2^{i+2}}{i!}$ for all $i > 4$. Thus $\frac{\partial \bar{z}}{\partial \mu}\big|_{\mu=0} > 0$ which concludes the proof of the first part of Proposition 1.

Now let's show that $\frac{\partial \hat{z}}{\partial \mu}\big|_{\mu=0} > \frac{\partial \bar{z}}{\partial \mu}\big|_{\mu=0}$. Using expressions for these derivatives and the same substitution ($x := \alpha \bar{z}_0$) this amounts to showing:

$$\frac{2(x^2w_2 + xw_1 - w_1^2)(xw_2 - w_1) - (4x^2 + xw_1w_2 - 2w_1^2)(2x - w_1)}{2\rho w_1(2x - w_1)(xw_2 - w_1)} > 0$$

Note that denominator is negative since $2x - w_1 < 0$ (trivial) and $xw_2 - w_1 > 0$ (shown above). Thus it remains to show that numerator (Num) is also negative. Opening the brackets, collecting terms and dividing by x yields:

$$Num = 2x^2w_2^2 - w_1^2w_2 + 2w_1^2 - 8x^2 - 2xw_1w_2 + 4xw_1$$

Plugging expressions from Appendix A.2.1:

$$\begin{aligned}
Num = {} & 2x^2\left(4 + 2\sum_{i=2,4,6...}^{\infty}\frac{2^ix^i}{i!}\right) - \left(2 + 2\sum_{i=2,4,6...}^{\infty}\frac{3^ix^i}{i!}\right) + \left(2 + 2\sum_{i=2,4,6...}^{\infty}\frac{x^i}{i!}\right) \\
& + 4\sum_{i=2,4,6...}^{\infty}\frac{2^ix^i}{i!} - 8x^2 - 4x\sum_{i=1,3,5...}^{\infty}\frac{2^ix^i}{i!} + 8x\sum_{i=1,3,5...}^{\infty}\frac{x^i}{i!} \\
= {} & 8x^2 + \frac{16}{2}x^4 + \sum_{i=6,8,10...}^{\infty}\frac{2^ix^i}{(i-2)!} - 2\frac{9}{2}x^2 - 2\frac{81}{24}x^4 - 2\sum_{i=6,8,10...}^{\infty}\frac{3^ix^i}{i!} \\
& + 2\frac{1}{2}x^2 + 2\frac{1}{24}x^4 + 2\sum_{i=6,8,10...}^{\infty}\frac{x^i}{i!} + 4\frac{4}{2}x^2 + 4\frac{16}{24}x^4 + 4\sum_{i=6,8,10...}^{\infty}\frac{2^ix^i}{i!} - 8x^2 \\
& - 2\frac{4}{1}x^2 - 2\frac{16}{6}x^4 - 2\sum_{i=6,8,10...}^{\infty}\frac{2^ix^i}{(i-1)!} + 8x^2 + 8\frac{1}{6}x^4 + 8\sum_{i=6,8,10...}^{\infty}\frac{x^i}{(i-1)!} \\
= {} & \sum_{i=6,8,10...}^{\infty}\left[\frac{2^i}{(i-2)!} - 2\frac{3^i}{i!} + \frac{2}{i!} + \frac{2^{i+2}}{i!} - \frac{2^{i+1}}{(i-1)!} + \frac{8}{(i-1)!}\right]x^i
\end{aligned}$$

Thus if $\frac{2}{i!}\left(2^{i-1}i(i-1) - 3^i + 1 + 2^{i+1} - 2^ii + 4i\right) \leq 0$ for all $i = \{6, 8, 10, \dots\}$ and inequality is strict for some i, then $Num < 0$. Define:

$$q(i) = 2^{i-1}i(i-1) - 3^i + 1 + 2^{i+1} - 2^ii + 4i$$

and note that $q(6) = 0$ and $q(8) < 0$. Now split $q(i)$ into two parts:

$$\begin{aligned}
q_1(i) &= 2^{i-1}i(i-1) + 2^{i+1} - 3^i \\
q_2(i) &= 1 + 4i - 2^ii \\
q(i) &= q_1(i) + q_2(i)
\end{aligned}$$

It is easy to see that $q_2(i) < 0$ for all $i > 3$. Let's show by induction that $q_1(i)$ is negative for all $i \geq 10$. First note that $q_1(10) < 0$. Now assume $q_1(i) < 0$. Rearranging, this implies:

$$i(i-1) < 2\left[\left(\frac{3}{2}\right)^i - 2\right]$$

Then for $i+1$ it holds:

$$(i+1)i = i(i-1)\frac{i+1}{i-1} < 2\left[\left(\frac{3}{2}\right)^i - 2\right]\frac{i+1}{i-1} < 2\left[\left(\frac{3}{2}\right)^{i+1} - 2\right]$$

where the first inequality is due to induction assumption and the last one is true for all $i > 5$. Thus $q_1(i) < 0$ for all $i \geq 10$ and same applies to $q(i)$, which concludes the proof.

A.2.3 Lemma 1

Let $\rho, \kappa, \sigma > 0$. Let $\Delta^+(\mu) = \hat{z}(\mu) - \underline{z}(\mu)$. Then:

$$\frac{\partial \Delta^+(0)}{\partial \mu} < \frac{1}{10}\frac{\overline{z}_0^2}{\sigma^2}$$

Proof. Using expressions for $\frac{\partial \underline{z}(0)}{\partial \mu}$ and $\frac{\partial \hat{z}(0)}{\partial \mu}$, and denoting $x := \alpha \overline{z}_0$, the above expressions can be written as:

$$\frac{10(4x^2+xw_1w_2-2w_1^2)(2x-w_1)-20(x^2w_2+xw_1-w_1^2)(xw_2-w_1)+w_1(xw_2-w_1)(2x-w_1)x^2}{20\rho w_1(2x-w_1)(xw_2-w_1)} > 0$$

Given that denominator is negative (as shown in proof of Proposition 1), it is required to show that numerator is positive. Opening the brackets, collecting terms and dividing by x, gives that numerator (Num) is negative if:

$$Num = 80x^2 - 20w_1^2 + 10w_1^2w_2 + x(20w_1w_2 - 40w_1 + w_1^3) - x^2(20w_2^2 + 2w_1^2 + w_1^2w_2) + 2x^3w_1w_2 < 0$$

Using expressions from Appendix A.2.1:

$$
\begin{aligned}
Num = {} & 80x^2 - 40\sum_{i=2,4,6...}^{\infty}\frac{2^i x^i}{i!} + 20\sum_{i=2,4,6...}^{\infty}\frac{3^i x^i}{i!} - 20\sum_{i=2,4,6...}^{\infty}\frac{x^i}{i!} \\
& + x\left(40\sum_{i=1,3,5...}^{\infty}\frac{2^i x^i}{i!} + 2\sum_{i=1,3,5...}^{\infty}\frac{3^i x^i}{i!} - 86\sum_{i=1,3,5...}^{\infty}\frac{x^i}{i!}\right) \\
& - x^2\left(80 + 44\sum_{i=2,4,6...}^{\infty}\frac{2^i x^i}{i!} + 2\sum_{i=2,4,6...}^{\infty}\frac{3^i x^i}{i!} - 2\sum_{i=2,4,6...}^{\infty}\frac{x^i}{i!}\right) + 4x^3\sum_{i=1,3,5...}^{\infty}\frac{2^i x^i}{i!} \\
& = 80x^2 - 40\frac{4}{2}x^2 - 40\sum_{i=4,6,8...}^{\infty}\frac{2^i x^i}{i!} + 20\frac{9}{2}x^2 + 20\sum_{i=4,6,8...}^{\infty}\frac{3^i x^i}{i!} - 20\frac{1}{2}x^2 - 20\sum_{i=4,6,8...}^{\infty}\frac{x^i}{i!} \\
& + 40\frac{2}{1}x^2 + 20\sum_{i=4,6,8...}^{\infty}\frac{2^i x^i}{(i-1)!} + 2\frac{3}{1}x^2 + 2\sum_{i=4,6,8...}^{\infty}\frac{3^{i-1}x^i}{(i-1)!} - 86x^2 - 86\sum_{i=4,6,8...}^{\infty}\frac{x^i}{(i-1)!} \\
& - 80x^2 - 11\sum_{i=4,6,8...}^{\infty}\frac{2^i x^i}{(i-2)!} - 2\sum_{i=4,6,8...}^{\infty}\frac{3^{i-2}x^i}{(i-2)!} + 2\sum_{i=4,6,8...}^{\infty}\frac{x^i}{(i-2)!} + 2\sum_{i=4,6,8...}^{\infty}\frac{2^{i-2}x^i}{(i-3)!} \\
& = \sum_{i=4,6,8...}^{\infty}\frac{2x^i}{i!}q(i)
\end{aligned}
$$

where

$$q(i) = 10(3^i - 2^{i+1} - 1) + i(10 \cdot 2^i + 3^{i-1} - 43) + i(i-1)(1 - 11 \cdot 2^{i-1} - 3^{i-2}) + i(i-1)(i-2)2^{i-2}$$

Thus if $q(i) \leq 0$ for all $i \in \{4, 6, 8...\}$ and $q(i) < 0$ for some $i \in \{4, 6, 8...\}$, then $Num < 0$ and Lemma 1 is proven. A direct computation gives that $q(4) = q(6) = q(8) = 0$ and $q(10) < 0$. Let's show that $q(i) < 0$ for all $i \geq 12$. Note that $q(i) < 0$ if and only if:

$$\underbrace{10(3^i - 2^{i+1} - 1)}_{q_1(i)} + \underbrace{i(10 \cdot 2^i + 3^{i-1} - 43)}_{q_2(i)} + \underbrace{i(i-1)(i-2)2^{i-2}}_{q_3(i)} < \underbrace{i(i-1)(3^{i-2} + 11 \cdot 2^{i-1} - 1)}_{q_4(i)}$$

Let's establish relations between these terms:

- $q_1(i) < \frac{1}{2}q_4(i)$:

$$
\begin{aligned}
q_1(i) < \frac{1}{2}q_4(i) & \iff 20(3^i - 2^{i+1} - 1) < i(i-1)(3^{i-2} + 11 \cdot 2^{i-1} - 1) \\
& \iff \underbrace{3^{i-2}(180 - i(i-1)) + i(i-1)}_{<0 \text{ for } i \geq 12} \underbrace{-2^{i-1}(80 + 11i(i-1)) - 20}_{<0} < 0
\end{aligned}
$$

Term in second bracket is trivially negative. To see why term in the first bracket is negative as well, consider a proof by induction. If $i = 12$, then $3^{i-2}(180 - i(i-1)) + i(i-1) < 0$. Suppose now that for some i, $3^{i-2}(180 - i(i-1)) < -i(i-1)$. Consider $i+1$:

$$\begin{aligned} 3^{i-1}(180 - (i+1)i) &< 3^{i-1}(180 - i(i-1)) = 3 \cdot 3^{i-2}(180 - i(i-1)) \\ &< \underbrace{-3i(i-1) < -i(i+1)}_{\text{for } i>2} \end{aligned}$$

where the second line follows from the induction assumption and the last one inequality is true for all $i > 2$. As a result, $q_1(i) < \frac{1}{2}q_4(i)$ for all $i \geq 12$.

- $q_2(i) < \frac{1}{4}q_4(i)$:

$$\begin{aligned} q_2(i) < \frac{1}{4}q_4(i) &\iff 4(10 \cdot 2^i + 3^{i-1} - 43) < (i-1)(3^{i-2} + 11 \cdot 2^{i-1} - 1) \\ &\iff \underbrace{3^{i-2}(13-i) + i}_{<0 \text{ for } i\geq 14} < \underbrace{2^{i-1}(11i - 91)}_{>0 \text{ for } i\geq 9} + 173 \end{aligned}$$

The right hand side is trivially positive for $i \geq 9$. To see why term on the left hand side is negative, consider $i \geq 14$ and rewrite it as:

$$3^{i-2}(13-i) + i < 0 \iff 3^{i-2} > \frac{i}{i-13}$$

Here, $\frac{i}{i-13}$ is a decreasing function of i, whereas 3^{i-2} is increasing. In addition, the inequality is true for $i = 14$ and thus it is true for all $i \geq 14$. Finally, direct computation shows that $q_2(i) < \frac{1}{4}q_4(i)$ for $i = 12$ and, as a result, $q_2(i) < \frac{1}{4}q_4(i)$ for all $i \geq 12$.

- $q_3(i) < \frac{1}{5}q_4(i)$:

$$q_3(i) < \frac{1}{5}q_4(i) \iff 2^{i-2}(5i - 22) < 3^{i-2} - 1$$

It suffices to show that $2^{i-2}5i < 3^{i-2} - 1$, which can be proven by induction. First, it holds for $i = 14$. Now assume that it holds for some i and consider $i+1$:

$$2^{i-1}5(i+1) = 2^{i-2}5i\frac{2(i+1)}{i} < \underbrace{(3^{i-2} - 1)\frac{2(i+1)}{i} < 3^{i-1} - 1}_{\text{for } i\geq 2}$$

where the first inequality follows from induction assumption and the second one can be seen by multiplying both sides with i and collecting terms, so that it is equivalent to $3^{i-2}(2-i) < i+2$ which holds trivially if $i \geq 2$. Finally, direct computation shows that $q_3(i) < \frac{1}{5}q_4(i)$ for $i = 12$ and so $q_3(i) < \frac{1}{5}q_4(i)$ for all $i \geq 12$.

It follows that $q_1(i) + q_2(i) + q_3(i) < \frac{19}{20}q_4(i) < q_4(i)$ for all $i \geq 12$ and thus $q(i) < 0$ for all $i \geq 8$, which concludes the proof.

A.2.4 Lemma 2

Let $\rho, \kappa, \sigma > 0$. If μ is small and non-zero, then the average price gap in the steady state $\bar{x}(\mu)$ is not equal to zero.

Proof. Due to the symmetry of the stationary distribution under zero drift, $\bar{x}(0) = 0$. It thus suffices to show that the derivative $\frac{\partial \bar{x}(0)}{\partial \mu}$ is not equal to zero. Recall from Appendix A.1.6 that:

$$\begin{aligned}
\frac{\partial \bar{x}(0)}{\partial \mu} &= \frac{2}{3}\frac{\partial \overline{z}(0)}{\partial \mu} + \frac{1}{3}\frac{\partial \hat{z}(0)}{\partial \mu} - \frac{\overline{z}_0^2}{6\sigma^2} \\
&= \frac{\partial \overline{z}(0)}{\partial \mu} + \frac{1}{3}\frac{\partial \Delta^+(0)}{\partial \mu} - \frac{\overline{z}_0^2}{6\sigma^2}
\end{aligned}$$

where $\Delta^+(\mu) = \hat{z}(\mu) - \underline{z}(\mu)$. From Lemma 1 it follows:

$$\begin{aligned}
\frac{\partial \bar{x}(0)}{\partial \mu} &< \frac{\partial \overline{z}(0)}{\partial \mu} + \frac{1}{30}\frac{\overline{z}_0^2}{\sigma^2} - \frac{\overline{z}_0^2}{6\sigma^2} \\
&= \frac{\partial \overline{z}(0)}{\partial \mu} - \frac{2}{15}\frac{\overline{z}_0^2}{\sigma^2}
\end{aligned}$$

Let me now show that $\frac{\partial \overline{z}(0)}{\partial \mu} - \frac{2}{15}\frac{\overline{z}_0^2}{\sigma^2} < 0$. Using the expression for $\frac{\partial \overline{z}(0)}{\partial \mu}$, rearranging terms and denoting $x := \alpha \overline{z}_0$, it is equivalent to showing that:

$$\frac{60x^2 + 15xw_1w_2 - 30w_1^2 - 2x^3w_1w_2 + 2x^2w_1^2}{30\rho(xw_1w_2 - w_1^2)} < 0$$

Note that the denominator is positive, as shown in the proof of Proposition 1. It thus suffices to show that the numerator is negative:

$$Num = 60x^2 + 15xw_1w_2 - 30w_1^2 - 2x^3w_1w_2 + 2x^2w_1^2 < 0$$

Using the expansion formulas from Appendix A.2.1, rewrite the numerator as:

$$
\begin{aligned}
Num &= 60x^2 + 30x \sum_{i=1,3,5...}^{\infty} \frac{2^i x^i}{i!} - 60 \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{i!} - 4x^3 \sum_{i=1,3,5...}^{\infty} \frac{2^i x^i}{i!} + 4x^2 \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{i!} \\
&= 60x^2 + 15 \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{(i-1)!} - 60 \sum_{i=2,4,6...}^{\infty} \frac{2^i x^i}{i!} - \sum_{i=4,6,8...}^{\infty} \frac{2^{i-1} x^i}{(i-3)!} + \sum_{i=4,6,8...}^{\infty} \frac{2^i x^i}{(i-2)!} \\
&= 15 \sum_{i=4,6,8...}^{\infty} \frac{2^i x^i}{(i-1)!} - 60 \sum_{i=4,6,8...}^{\infty} \frac{2^i x^i}{i!} - \sum_{i=4,6,8...}^{\infty} \frac{2^{i-1} x^i}{(i-3)!} + \sum_{i=4,6,8...}^{\infty} \frac{2^i x^i}{(i-2)!} \\
&= \sum_{i=4,6,8...}^{\infty} \frac{2^{i-1} x^i}{i!} \Big[\underbrace{30i - 120 - i(i-1)(i-2) + 2i(i-1)}_{q(i)} \Big]
\end{aligned}
$$

If $q(i) \leq 0$ for all $i \in \{4, 6, 8...\}$ and $q(i) < 0$ for some of these i, it would follow that $Num < 0$ and Lemma 2 is proven.

Note first that $q(4) = q(6) = 0$, whereas $q(8) < 0$ and $q(10) < 0$. Let me prove by induction that $q(i) + 120 < 0$ for any $i \geq 10$. Suppose that for some i, $q(i) + 120 < 0$. Consider $i + 1$:

$$
\begin{gathered}
q(i+1) + 120 = 30(i+1) - (i+1)i(i-1) + 2(i+1)i < 0 \iff \\
30 - i(i-1) + 2i < 0 \iff 30 - i(i-3) < 0
\end{gathered}
$$

The last inequality is trivially satisfied for any $i \geq 10$, which concludes the proof.

A.2.5 Proof of Proposition 2

First, for convenience, denote $\hat{\Theta}(\delta) = \frac{\partial \Theta(\delta,\mu)}{\partial \mu}\big|_{\mu=0}$. Consider $\delta > 0$. Note that:

$$
\hat{\Theta}(\delta) = \begin{cases} 0, & \text{for } \delta = 0 \\ \frac{\overline{z}_0^2}{4\sigma^2} - \frac{1}{6}\frac{\partial \Delta^+(0)}{\partial \mu}, & \text{for } \delta = \overline{z}_0 \\ \frac{\overline{z}_0^2}{6\sigma^2} + \frac{2}{3}\frac{\partial \Delta^+(0)}{\partial \mu}, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}
$$

And thus $\hat{\Theta}(\overline{z}_0) > 0$ by Lemma 1, and $\hat{\Theta}(\delta) > 0$ for all $\delta \geq 2\overline{z}_0$ since $\frac{\partial \Delta^+(0)}{\partial \mu} > 0$ by Proposition 1.

Consider now $\delta \in (0, \overline{z}_0)$. For such δ:

$$\begin{aligned}\hat{\Theta}'(\delta) &= \delta\left[\frac{6\overline{z}_0^2 - 2\delta^2 - 3\delta\overline{z}_0}{6\sigma^2\overline{z}_0^2} - \frac{\delta}{2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu}\right] > \delta\left[\frac{6\overline{z}_0^2 - 2\overline{z}_0^2 - 3\overline{z}_0^2}{6\sigma^2\overline{z}_0^2} - \frac{\overline{z}_0}{2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu}\right] \\ &= \delta\left[\frac{1}{6\sigma^2} - \frac{1}{2\overline{z}_0^2}\frac{\partial\Delta^+(0)}{\partial\mu}\right] > 0\end{aligned}$$

where first inequality is due to $\delta < \overline{z}_0$ and second one due to Lemma 1. It follows that $\hat{\Theta}(\delta)$ is strictly increasing over $(0, \overline{z}_0)$ and since $\hat{\Theta}(0) = 0$ it follows that $\hat{\Theta}(\delta) > 0$ for all $\delta \in (0, \overline{z}_0]$.

Consider now $\delta \in (\overline{z}_0, 2\overline{z}_0)$. For such δ:

$$\hat{\Theta}'(\delta) = \frac{2\delta^3 - 3\delta^2\overline{z}_0 - 6\delta\overline{z}_0^2 + 8\overline{z}_0^3}{6\sigma^2\overline{z}_0^2} - \frac{\delta^2 - 4\overline{z}_0^2}{2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu}$$

so that $\lim_{\delta\downarrow\overline{z}_0}\hat{\Theta}'(\delta) = \frac{\overline{z}_0}{6\sigma^2} + \frac{3}{2\overline{z}_0}\frac{\partial\Delta^+(0)}{\partial\mu} > 0$. Given that $\hat{\Theta}(\overline{z}_0), \hat{\Theta}(2\overline{z}_0) > 0$, the only case when $\hat{\Theta}(\delta)$ is negative for some $\delta \in (\overline{z}_0, 2\overline{z}_0)$ is if its derivative $\hat{\Theta}'(\delta)$ becomes negative and then again positive, i.e. switches its sign at least twice. To see if that is the case, consider second and third derivatives:

$$\begin{aligned}\hat{\Theta}''(\delta) &= \frac{\delta^2 - \delta\overline{z}_0 - \overline{z}_0^2}{\sigma^2\overline{z}_0^2} - \frac{\delta}{\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} \\ \hat{\Theta}'''(\delta) &= \frac{2\delta - \overline{z}_0}{\sigma^2\overline{z}_0^2} - \frac{1}{\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} > \frac{\overline{z}_0}{\sigma^2\overline{z}_0^2} - \frac{1}{\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} > 0\end{aligned}$$

where first inequality follows since $\delta > \overline{z}_0$ and second one from Lemma 1. Third derivative is strictly positive for all $\delta \in (\overline{z}_0, 2\overline{z}_0)$ and thus second derivative is monotonic and can only cross zero at most once. It follows that first derivative $\hat{\Theta}'(\delta)$ can switch its sign at most once and thus $\hat{\Theta}(\delta)$ is strictly positive for all $\delta \in (\overline{z}_0, 2\overline{z}_0)$. Given previous results, it follows that $\frac{\Theta(\delta,\mu)}{\partial\mu}\Big|_{\mu=0} > 0$ for all $\delta > 0$. Noting that impact effect is symmetric around zero drift ($\Theta(-\delta, \mu) = -\Theta(\delta, -\mu)$) provides that $\frac{\Theta(-\delta,0)}{\partial\mu} = \frac{\Theta(\delta,0)}{\partial\mu} > 0$ which concludes the proof.

A.2.6 Proof of Proposition 3

Let $\hat{m}(z,t)$ denote the expected cumulative deviation of g from its steady state until time t, conditional on initial value $z(0) = z$:

$$\hat{m}(z,t) = \mathbb{E}\left(\int_0^t (g(z(s)) - \bar{g})ds \,\middle|\, z(0) = z\right)$$

Denote $\hat{m}(z) = \lim_{t\to\infty} \hat{m}(z,t)$ and thus:

$$CIRF(F_0) = \int_{\underline{z}}^{\overline{z}} \hat{m}(z) dF_0(z)$$

Let τ_i be the i-th adjustment and let $t_a \wedge t_b = min\{t_a, t_b\}$. Fix a staring value z and consider the cumulated deviation of g form its steady state until $t > 0$, writing all the random variables explicitly as a function of the underlying outcome ω:

$$\begin{aligned}\int_0^t (g(z(s,\omega)) - \bar{g})ds = &\int_0^{\tau_1(\omega)\wedge t} (g(z(s,\omega)) - \bar{g})ds + \sum_{i=1}^{N-1} \int_{\tau_i(\omega)\wedge t}^{\tau_{i+1}(\omega)\wedge t} (g(z(s,\omega)) - \bar{g})ds \\ &+ \int_{\tau_N(\omega)\wedge t}^{t} (g(z(s,\omega)) - \bar{g})ds\end{aligned}$$

for some fixed $N \geq 1$. Take the limit of the above expression as $N \to \infty$. For a fixed horizon t and outcome ω there will be $n(t,\omega)$ adjustments between time 0 and t. Let $N(t,\omega) = \max\{1, n(t,\omega)\}$. Then:

$$\begin{aligned}\int_0^t (g(z(s,\omega)) - \bar{g})ds = &\int_0^{\tau_1(\omega)\wedge t} (g(z(s,\omega)) - \bar{g})ds + \sum_{i=1}^{N(t,\omega)-1} \int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} (g(z(s,\omega)) - \bar{g})ds \\ &+ \int_{\tau_{N(t,\omega)}(\omega)\wedge t}^{t} (g(z(s,\omega)) - \bar{g})ds\end{aligned}$$

Applying conditional expectation ($\mathbb{E}_z(\cdot) = \mathbb{E}(\cdot|z(0,\omega) = z)$) yields an expression for $\hat{m}(z,t)$:

$$\begin{aligned}\hat{m}(z,t) = &\mathbb{E}_z\left(\int_0^{\tau_1(\omega)\wedge t} (g(z(s,\omega)) - \bar{g})ds\right) \\ &+ \sum_{i=1}^{\infty} \mathbb{E}_z\left(\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} (g(z(s,\omega)) - \bar{g})ds \middle| N(t,\omega) \geq i+1\right) \mathbb{P}_z\left(N(t,\omega) \geq i+1\right) \\ &+ \mathbb{E}_z\left(\int_{\tau_{N(t,\omega)}(\omega)\wedge t}^{t} (g(z(s,\omega)) - \bar{g})ds\right)\end{aligned}$$

Where $\mathbb{P}_z\left(N(t,\omega) \geq i+1\right)$ is the probability that number of adjustments until t exceeds $i+1$ conditional on $z(0,\omega) = z$. Note that once we take expectation with respect to ω, the finite sum from the previous expression becomes infinite. That is due to the fact that for any $t > 0$ and any M there exists ω such that $N(t,\omega) > M$, which follows from the fact that increments of $z(t)$ are normally distributed. Each summand i is the expected cumulated deviation between i-th and $(i+1)$-th adjustment, conditional on there being at least $i+1$ adjustments, and weighted with corresponding probability.

Finally, take the limit as $t \to \infty$. For every z and every $i \in R_+$, $\mathbb{P}_z\left(N(t,\omega) \geq i+1\right)$ converges to one and the conditional expectation in second line converges to unconditional one. Also $N(t,\omega)$ converges to $n(t,\omega)$, $\tau_1(\omega) \wedge t$ converges to $\tau_1(\omega)$ and $\tau_{N(t,\omega)}(\omega) \wedge t \to \tau_{n(t,\omega)}(\omega)$. As has been shown in Baley and Blanco (2020), $\mathbb{E}_z\left(\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} (g(z(s,\omega)) - \bar{g})ds\right) = 0$ for all i, and thus:

$$\hat{m}(z) = \lim_{t\to\infty} \hat{m}(z,t) = m(z) + \tilde{m}(z)$$

where

$$m(z) = \mathbb{E}\left(\int_0^{\tau_1(\omega)} (g(z(s,\omega)) - \bar{g})ds \middle| z(0,\omega) = z\right)$$

$$\tilde{m}(z) = \lim_{t\to\infty} \mathbb{E}\left(\int_{\tau_{n(t,\omega)}(\omega)}^{t} (g(z(s,\omega)) - \bar{g})ds \middle| z(0,\omega) = z\right)$$

Note that due to Markov property, $\tilde{m}(z)$ does not depend on z since after the first adjustment initial condition does not matter and expectation becomes unconditional, so that $\tilde{m}(z) = \tilde{m} = \lim_{t\to\infty} \mathbb{E}\left(\int_{\tau_{n(t,\omega)}(\omega)}^{t} (g(z(s,\omega)) - \bar{g})ds\right)$. Thus:

$$CIRF(F_0) = \int_{\underline{z}}^{\overline{z}} m(z)dF_0(z) + \tilde{m}$$

which concludes the proof.

A.2.7 Proof of Proposition 4

Let $\hat{m}(r,z,t)$ denote the expected discounted cumulative deviation of g from its steady state until time t, conditional on initial value $z(0) = z$:

$$\hat{m}(r,z,t) = \mathbb{E}\left(\int_0^{t} e^{-rs}(g(z(s)) - \bar{g})ds \middle| z(0) = z\right)$$

with $r > 0$. Denote $\hat{m}(r,z) = \lim_{t\to\infty} \hat{m}(r.z,t)$ so that discounted cumulative impulse response is given by:

$$DCIRF(r,F_0) = \int_{\underline{z}}^{\overline{z}} \hat{m}(r,z)dF_0(z)$$

Let τ_i be the i-th adjustment and let $t_a \wedge t_b = min\{t_a, t_b\}$. Fix a staring value z and consider the discounted cumulated deviation of g form its steady state until $t > 0$, writing all the random variables explicitly as a function of the underlying outcome ω:

$$\begin{aligned}\int_0^t e^{-rs}(g(z(s,\omega))-\bar{g})ds = &\int_0^{\tau_1(\omega)\wedge t} e^{-rs}(g(z(s,\omega))-\bar{g})ds + \\ &\sum_{i=1}^{N-1}\int_{\tau_i(\omega)\wedge t}^{\tau_{i+1}(\omega)\wedge t} e^{-rs}(g(z(s,\omega))-\bar{g})ds + \\ &\int_{\tau_N(\omega)\wedge t}^{t} e^{-rs}(g(z(s,\omega))-\bar{g})ds\end{aligned}$$

for some fixed $N \geq 1$. Take the limit of the above expression as $N \to \infty$. For a fixed horizon t and outcome ω there will be $n(t,\omega)$ adjustments between time 0 and t. Let $N(t,\omega) = \max\{1, n(t,\omega)\}$. Then:

$$\begin{aligned}\int_0^t e^{-rs}(g(z(s,\omega))-\bar{g})ds = &\int_0^{\tau_1(\omega)\wedge t} e^{-rs}(g(z(s,\omega))-\bar{g})ds + \\ &\sum_{i=1}^{N(t,\omega)-1}\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} e^{-rs}(g(z(s,\omega))-\bar{g})ds + \\ &\int_{\tau_{N(t,\omega)}(\omega)\wedge t}^{t} e^{-rs}(g(z(s,\omega))-\bar{g})ds\end{aligned}$$

Applying conditional expectation ($\mathbb{E}_z(\cdot) = \mathbb{E}(\cdot|z(0,\omega) = z)$) yields an expression for $\hat{m}(z,t)$:

$$\begin{aligned}\hat{m}(r,z,t) = &\mathbb{E}_z\left(\int_0^{\tau_1(\omega)\wedge t} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right) \\ &+ \sum_{i=1}^{\infty}\mathbb{E}_z\left(\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds \middle| N(t,\omega) \geq i+1\right)\mathbb{P}_z\left(N(t,\omega) \geq i+1\right) \\ &+ \mathbb{E}_z\left(\int_{\tau_{N(t,\omega)}(\omega)\wedge t}^{t} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right)\end{aligned}$$

Where $\mathbb{P}_z\left(N(t,\omega) \geq i+1\right)$ is the probability that number of adjustments until t exceeds $i+1$ conditional on $z(0,\omega) = z$. Note that once we take expectation with respect to ω, the finite sum from the previous expression becomes infinite. That is due to the fact that for any $t > 0$ and any M there exists ω such that $N(t,\omega) > M$, which follows from the fact that increments of $z(t)$ are normally distributed. Each summand i is the expected cumulated deviation between i-th and $(i+1)$-th adjustment, conditional on there being at least $i+1$ adjustments, and weighted with corresponding probability.

Finally, take the limit as $t \to \infty$. For every z and every $i \in R_+$, $\mathbb{P}_z\left(N(t,\omega) \geq i+1\right)$ converges to one and the conditional expectation in second line converges to unconditional one. Also $N(t,\omega)$ converges to $n(t,\omega)$, $\tau_1(\omega) \wedge t$ converges to $\tau_1(\omega)$ and $\tau_{N(t,\omega)}(\omega) \wedge t \to \tau_{n(t,\omega)}(\omega)$. Due to $r > 0$, the last summand converges to zero and thus:

$$\hat{m}(r,z) = \lim_{t\to\infty} \hat{m}(r,z,t) = \overbrace{\mathbb{E}_z\left(\int_0^{\tau(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right)}^{m(r,z)} + \sum_{i=1}^{\infty} \mathbb{E}_z\left(\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right)$$

Because of discounting, expected deviations between adjustments are not zero anymore. However one can still characterize them. First, consider some $i \geq 1$ and rewrite as:

$$\mathbb{E}_z\left(\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right) = \mathbb{E}_z\left(e^{-r\tau_i(\omega)}\int_0^{\tau_{i+1}(\omega)-\tau_i(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right)$$

Note that due to strong Markov property, expectation of the integral does not depend on i or z, whereas expectation of $e^{-\tau_i(\omega)}$ depends on both. Thus the two terms are independent and we can split the expectation:

$$\mathbb{E}_z\left(e^{-r\tau_i(\omega)}\int_0^{\tau_{i+1}(\omega)-\tau_i(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right) = \mathbb{E}_z\left(e^{-r\tau_i(\omega)}\right) \cdot \underbrace{\mathbb{E}_{\hat{z}}\left(\int_0^{\tau(\omega)} e^{-rs}(g(z(s,\omega)) - \bar{g})ds\right)}_{m(r,\hat{z})}$$

Now second term is the expectation of cumulated deviations until first adjustment conditional on starting at the return point: $z(0) = \hat{z}$. Denote the first term for $i=1$ by $q(r,z)=\mathbb{E}_z\left(e^{-r\tau(\omega)}\right)$. Then for any $i \geq 1$:

$$\begin{aligned}
\mathbb{E}_z\left(e^{-r\tau_i(\omega)}\right) &= \mathbb{E}_z\left(e^{-r\tau_1(\omega)}\cdot e^{-r(\tau_2(\omega)-\tau_1(\omega))}\cdots e^{-r(\tau_i(\omega)-\tau_{i-1}(\omega))}\right) \\
&= \mathbb{E}_z\left(e^{-r\tau_1(\omega)}\right)\cdot \mathbb{E}_z\left(e^{-r(\tau_2(\omega)-\tau_1(\omega))}\right)\cdots \mathbb{E}_z\left(e^{-r(\tau_i(\omega)-\tau_{i-1}(\omega))}\right) \\
&= \underbrace{\mathbb{E}_z\left(e^{-r\tau(\omega)}\right)}_{q(r,z)}\cdot\underbrace{\mathbb{E}_{\hat{z}}\left(e^{-r\tau(\omega)}\right)}_{q(r,\hat{z})}\cdots\underbrace{\mathbb{E}_{\hat{z}}\left(e^{-r\tau(\omega)}\right)}_{q(r,\hat{z})} \\
&= q(r,z)q(r,\hat{z})^{i-1}
\end{aligned}$$

where second and third lines follow due to strong Markov property of $z(t)$. Because of this property, times between adjustments are independent (2nd line) and initial condition $z(0)=z$ is irrelevant once there was an adjustment (3rd line). Thus:

$$\sum_{i=1}^{\infty}\mathbb{E}_z\left(\int_{\tau_i(\omega)}^{\tau_{i+1}(\omega)} e^{-rs}(g(z(s,\omega))-\bar{g})ds\right)=\sum_{i=1}^{\infty}q(r,z)q(r,\hat{z})^{i-1}m(r,\hat{z})=\frac{q(r,z)}{1-q(r,\hat{z})}m(r,\hat{z})$$

and so:

$$DCIRF(r,F_0)=\int_{\underline{z}}^{\overline{z}} m(r,z)dF_0(z)+\frac{m(r,\hat{z})}{1-q(r,\hat{z})}\int_{\underline{z}}^{\overline{z}} q(r,z)dF_0(z)$$

Now in order to obtain undiscounted CIRF, it remains to take the limit as $r\to 0$. Note that $\lim\limits_{r\to 0} m(r,z)=m(z)$ where $m(z)=\mathbb{E}\left(\int_0^{\tau}(g(z(s))-\bar{g})ds\middle|z(0)=z\right)$ and $\lim\limits_{r\to 0} q(r,z)=1$. This implies that second integral converges to 1. In addition, since $m(\hat{z})=0$, as shown in Baley and Blanco (2020), $\lim\limits_{r\to 0} m(r,\hat{z})=0$, and the coefficient in front of the second integral converges to some finite number. We can further simplify the expression by noting that:

$$q(r,z)=\mathbb{E}_z\left(e^{-r\tau(\omega)}\right)=1-r\mathbb{E}_z\left(\int_0^{\tau(\omega)} e^{-rs}ds\right)$$

and since $\lim\limits_{r\to 0}\mathbb{E}_z\left(\int_0^{\tau(\omega)} e^{-rs}ds\right)=\mathbb{E}_z\left(\tau(\omega)\right)$, for small values of r, $1-q(r,\hat{z})$ behaves like $r\mathbb{E}_z\left(\tau(\omega)\right)$. Thus CIRF can be expressed as:

$$CIRF(F_0)=\lim_{r\to 0} DCIRF(r,F_0)=\int_{\underline{z}}^{\overline{z}} m(z)dF_0(z)+\frac{1}{\mathbb{E}\left(\tau(\omega)\middle|z(0,\omega)=\hat{z}\right)}\lim_{r\to 0}\frac{m(r,\hat{z})}{r}$$

where

$$m(z)=\mathbb{E}\left(\int_0^{\tau(\omega)}(g(z(s,\omega))-\bar{g})ds\ \middle|\ z(0,\omega)=z\right)$$

$$m(r,\hat{z})=\mathbb{E}\left(\int_0^{\tau(\omega)} e^{-rs}(g(z(s,\omega))-\bar{g})ds\ \middle|\ z(0,\omega)=\hat{z}\right)$$

which concludes the proof.

A.2.8 Proof of Proposition 5

First, for convenience, denote $\hat{M}(\delta) = \frac{\partial M(\delta,\mu)}{\partial \mu}\Big|_{\mu=0}$ and consider $\delta > 0$. Note that:

$$\hat{M}(\delta) = \begin{cases} 0, & \text{for } \delta = 0 \\ -\frac{17\overline{z}_0^4}{360\sigma^4} - \frac{13\overline{z}_0^2}{180\sigma^2}\frac{\partial \Delta^+(0)}{\partial \mu}, & \text{for } \delta = \overline{z}_0 \\ -\frac{\overline{z}_0^4}{60\sigma^4} - \frac{\overline{z}_0^2}{5\sigma^2}\frac{\partial \Delta^+(0)}{\partial \mu}, & \text{for } \delta \geq 2\overline{z}_0 \end{cases}$$

so that $\hat{M}(\overline{z}_0) < 0$ and $\hat{M}(\delta) < 0$ for all $\delta \geq 2\overline{z}_0$ since $\frac{\partial \Delta^+(0)}{\partial \mu} > 0$ by Proposition 1.

Now let's show that $\hat{M}(\delta) < 0$ for any $\delta > 0$. First, consider $\delta \in (0, \overline{z})$ and denote by $\hat{M}^k_-(\overline{z}_0)$ the limit of k-th derivative of $\hat{M}(\delta)$ for $\delta \uparrow \overline{z}_0$. The proof consists of five claims. To ease exposition, proofs of the claims are provided at the end of this section.

1a. $\hat{M}^V(0) < 0$ and $\hat{M}^V_-(\overline{z}) < 0$. In addition, $\hat{M}^V(\delta)$ is linear for $\delta \in (0, \overline{z}_0)$ and thus $\hat{M}^V(\delta) < 0$ for all $\delta \in (0, \overline{z}_0)$. This implies that $\hat{M}^{IV}(\delta)$ is strictly decreasing for $\delta \in (0, \overline{z}_0)$.

2a. $\hat{M}^{IV}(0) > 0$ and $\hat{M}^{IV}_-(\overline{z}) < 0$. Together with (1a) this implies that $\hat{M}^{IV}(\delta)$ crosses zeros once in $(0, \overline{z}_0)$ and thus $\hat{M}^{III}(\delta)$ is strictly concave and singe-peaked in $(0, \overline{z}_0)$.

3a. $\hat{M}^{III}(0) > 0$ and $\hat{M}^{III}_-(\overline{z}) < 0$. Together with (2a) it implies that $\hat{M}^{III}(\delta)$ crosses zeros once in $(0, \overline{z}_0)$ so that $\hat{M}^{II}(\delta)$ first increases and then decreases as δ goes from 0 to $\overline{z}_0$.

4a. $\hat{M}^{II}(0) < 0$ and $\hat{M}^{II}_-(\overline{z}) > 0$. Together with (3a) it implies that $\hat{M}^{II}(\delta)$ crosses zero once in $(0, \overline{z}_0)$ and thus $\hat{M}^{I}(\delta)$ first decreases and then increases as δ goes from 0 to $\overline{z}_0$.

5a. $\hat{M}^{I}(0) = 0$. Together with (4a) this implies that $\hat{M}^{I}(\delta)$ crosses zero *at most* once in $(0, \overline{z}_0)$.

Finally, since $\hat{M}(0) = 0$, $\hat{M}(\overline{z}_0) < 0$, $\hat{M}^{I}(0) = 0$, $\hat{M}^{II}(0) < 0$ and $\hat{M}^{I}(\delta)$ crosses zero at most once in $(0, \overline{z}_0)$, it follows that $\hat{M}(\delta) < 0$ for all $\delta \in (0, \overline{z}_0)$. In order to have $\hat{M}(\delta) \geq 0$ for some $\delta \in (0, \overline{z}_0)$, it must be the case that $\hat{M}^{I}(\delta)$

crosses zero at least twice, which contradicts (5a).

Now consider $\delta \in (\overline{z}_0, 2\overline{z}_0)$ and denote by $\hat{M}^k_+(\overline{z}_0)$ the limit of k-th derivative of $\hat{M}(\delta)$ for $\delta \downarrow \overline{z}_0$. The proof consists of five claims, proofs of which are also delegated to the end of this section.

1b. $\hat{M}^V_+(\overline{z}_0) > 0$ and $\hat{M}^V(2\overline{z}_0) > 0$. In addition, $\hat{M}^V(\delta)$ is linear for $\delta \in (\overline{z}_0, 2\overline{z}_0)$ and thus $\hat{M}^V(\delta) > 0$ for all $\delta \in (\overline{z}_0, 2\overline{z}_0)$. This implies that $\hat{M}^{IV}(\delta)$ is strictly increasing for $\delta \in (\overline{z}_0, 2\overline{z}_0)$.

2b. $\hat{M}^{IV}_+(\overline{z}_0) < 0$ and $\hat{M}^{IV}(2\overline{z}_0) > 0$. Together with (1b) this implies that $\hat{M}^{IV}(\delta)$ crosses zeros once in $(\overline{z}_0, 2\overline{z}_0)$ and thus $\hat{M}^{III}(\delta)$ is strictly convex in $(\overline{z}_0, 2\overline{z}_0)$.

3b. $\hat{M}^{III}_+(\overline{z}_0) < 0$ and $\hat{M}^{III}(2\overline{z}_0) > 0$. Together with (2b) it implies that $\hat{M}^{III}(\delta)$ crosses zeros once in $(\overline{z}_0, 2\overline{z}_0)$ so that $\hat{M}^{II}(\delta)$ first decreases and then increases as δ goes from $\overline{z}_0$ to $2\overline{z}_0$.

4b. $\hat{M}^{II}_+(\overline{z}_0) > 0$ and $\hat{M}^{II}(2\overline{z}_0) = 0$. Together with (3b) it implies that $\hat{M}^{II}(\delta)$ crosses zero once in $(\overline{z}_0, 2\overline{z}_0)$ and thus $\hat{M}^{I}(\delta)$ first increases and then decreases as δ goes from $\overline{z}_0$ to $2\overline{z}_0$.

5b. $\hat{M}^{I}(2\overline{z}_0) = 0$. Together with (4b) and (3b) this implies that $\hat{M}^{I}(\delta)$ crosses zero *at most* once in $\overline{z}_0$ to $2\overline{z}_0$.

Finally, since $\hat{M}(\overline{z}_0) < 0$, $\hat{M}(2\overline{z}_0) < 0$, $\hat{M}^{I}(2\overline{z}_0) = 0$, $\hat{M}^{II}(2\overline{z}_0) = 0$, $\hat{M}^{III}(2\overline{z}_0) > 0$ and $\hat{M}^{I}(\delta)$ crosses zero at most once in $(\overline{z}, 2\overline{z})$, it follows that $\hat{M}(\delta) < 0$ for all $\delta \in (\overline{z}, 2\overline{z})$. In order to have $\hat{M}(\delta) \geq 0$ for some $\delta \in (\overline{z}, 2\overline{z})$, it must be the case that $\hat{M}^{I}(\delta)$ crosses zero at least twice, which contradicts (5b).

Altogether, this implies that $\hat{M}(\delta) < 0$ for all $\delta > 0$. Note that $M(\delta, \mu)$ is symmetric in the sense that $M(-\delta, \mu) = -M(\delta, -\mu)$, so that $\frac{\partial M(-\delta,0)}{\partial \mu} = \frac{\partial M(\delta,0)}{\partial \mu} < 0$, which concludes the proof. Below I prove claims used above.

Consider $\delta > 0$. Since $\hat{M}(\delta)$ is a polynomial of degree 6, it follows that

$\hat{M}^V(\delta)$ is a linear function. Direct computation yields:

$$\begin{aligned}
\hat{M}^V(0) &= -\frac{6}{\sigma^4\overline{z}_0} - \frac{2}{\sigma^2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^V_-(\overline{z}_0) &= -\frac{14}{\sigma^4\overline{z}_0} - \frac{2}{\sigma^2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^V_+(\overline{z}_0) &= \frac{2}{\sigma^4\overline{z}_0} - \frac{2}{\sigma^2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^V(2\overline{z}_0) &= \frac{10}{\sigma^4\overline{z}_0} - \frac{2}{\sigma^2\overline{z}_0^3}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{IV}(0) &= \frac{3}{\sigma^4} - \frac{4}{3\sigma^2\overline{z}_0^2}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{IV}_-(\overline{z}_0) &= -\frac{7}{\sigma^4} - \frac{10}{3\sigma^2\overline{z}_0^2}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{IV}_+(\overline{z}_0) &= -\frac{1}{\sigma^4} - \frac{2}{3\sigma^2\overline{z}_0^2}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{IV}(2\overline{z}_0) &= \frac{5}{\sigma^4} - \frac{8}{3\sigma^2\overline{z}_0^2}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{III}(0) &= \frac{\overline{z}_0}{3\sigma^4} \\
\hat{M}^{III}_-(\overline{z}_0) &= -\frac{\overline{z}_0}{\sigma^4} - \frac{7}{3\sigma^2\overline{z}_0}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{III}_+(\overline{z}_0) &= -\frac{\overline{z}_0}{3\sigma^4} + \frac{1}{3\sigma^2\overline{z}_0}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{III}(2\overline{z}_0) &= \frac{\overline{z}_0}{\sigma^4} - \frac{4}{3\sigma^2\overline{z}_0}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{II}(0) &= -\frac{\overline{z}_0^2}{3\sigma^4} \\
\hat{M}^{II}_-(\overline{z}_0) &= \frac{\overline{z}_0^2}{6\sigma^4} - \frac{1}{\sigma^2}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{II}_+(\overline{z}_0) &= \frac{\overline{z}_0^2}{6\sigma^4} + \frac{1}{3\sigma^2}\frac{\partial\Delta^+(0)}{\partial\mu} \\
\hat{M}^{II}(2\overline{z}_0) &= 0 \\
\hat{M}^{I}(0) &= 0 \\
\hat{M}^{I}(2\overline{z}_0) &= 0
\end{aligned}$$

Inequalities in (1a - 5a) and (1b - 5b) follow either trivially, or due to Proposition 1 or due to Lemma 1.

A.2.9 Proof of Proposition 6

Let $\mu > 0$ and small. First, consider impact effect $\Theta(\delta, \mu)$. Its first order approximation with respect to drift is given by:

$$\Theta(\delta, \mu) = \Theta(\delta, 0) + \frac{\partial \Theta(\delta, 0)}{\partial \mu}\mu$$

Since for $\delta \geq 2\overline{z}_0$, $\Theta(\delta, 0) = \delta$ and $\frac{\partial \Theta(\delta,0)}{\partial \mu} > 0$ by Proposition 2, it follows that:

$$\Theta(\delta, \mu) - \delta > 0 \quad \text{for} \quad \delta \geq 2\overline{z}_0$$

Since both $\Theta(\delta, 0)$ and $\frac{\partial \Theta(\delta,0)}{\partial \mu}$ are second order in δ for small shocks, it follows that:

$$\Theta(\delta, \mu) - \delta < 0 \quad \text{for some small } \delta > 0$$

Thus, due to continuity of $\Theta(\delta, \mu)$, there exists $\delta_\Theta(\mu) \in (0, 2\overline{z}_0)$ such that $\Theta(\delta_\Theta(\mu), \mu) - \delta = 0$ and $\Theta(\delta, \mu) - \delta > 0$ for all $\delta > \delta_\Theta(\mu)$. Finally, since width of inaction region $\overline{z}(\mu) - \underline{z}(\mu)$ does not change with μ to first order (Proposition 1), and $\delta_\Theta(\mu) < 2\overline{z}_0$, it follows that $\delta_\Theta(\mu) < \overline{z}(\mu) - \underline{z}(\mu)$ if μ is sufficiently small.

Now, consider cumulative response $M(\delta, \mu)$. Its first order approximation with respect to drift is given by:

$$M(\delta, \mu) = M(\delta, 0) + \frac{\partial M(\delta, 0)}{\partial \mu}\mu$$

Since for $\delta \geq 2\overline{z}_0$, $M(\delta, 0) = 0$ and $\frac{\partial M(\delta,0)}{\partial \mu} < 0$ by Proposition 5, it follows that:

$$M(\delta, \mu) < 0 \quad \text{for} \quad \delta \geq 2\overline{z}_0$$

Since $M(\delta, 0)$ is first order and $\frac{\partial M(\delta,0)}{\partial \mu}$ is second order in δ for small shocks, it follows that:

$$M(\delta, \mu) > 0 \quad \text{for some small } \delta > 0$$

Thus, due to continuity of $M(\delta, \mu)$, there exists $\delta_M(\mu) \in (0, 2\overline{z}_0)$ such that $M(\delta_M(\mu), \mu) = 0$ and $M(\delta, \mu) < 0$ for all $\delta > \delta_M(\mu)$. Similar logic as before leads to $\delta_M(\mu) < \overline{z}(\mu) - \underline{z}(\mu)$ if μ is sufficiently small.

A.2.10 Lemma 3

Let $\rho, \kappa, \sigma > 0$. Then the CIR for a shock to the drift $\frac{\partial M(0)}{\partial \mu}$ (see Appendix A.1.7) is positive:

$$\frac{\partial M(0)}{\partial \mu} = \frac{\overline{z}_0^2}{180\sigma^2}\left[\frac{6\overline{z}_0^2}{\sigma^2} - 16\frac{\partial \underline{z}(0)}{\partial \mu} - 14\frac{\partial \hat{z}(0)}{\partial \mu}\right] > 0$$

Proof. It suffices to show that $\frac{3\overline{z}_0^2}{8\sigma^2} - \left(\frac{\partial \underline{z}(0)}{\partial \mu} + \frac{\partial \hat{z}(0)}{\partial \mu}\right) > 0$. Using expression from Appendix A.1.3, letting $x := \alpha \overline{z}(0)$ and collecting terms yields:

$$\frac{3\overline{z}_0^2}{8\sigma^2} - \left(\frac{\partial \underline{z}(0)}{\partial \mu} + \frac{\partial \hat{z}(0)}{\partial \mu}\right) =$$
$$-\frac{32w_1^3 - x(48w_1^2 + 24w_1^2 w_2) + x^2(16w_1w_2 - 32w_1 - 3w_1^3) + x^3(64 + 16w_2^2 + 6w_1^2 + 3w_1^2 w_2) - 6x^4 w_1 w_2}{16\rho(xw_1w_2 - w_1^2)(2x - w_1)}$$

As shown previously, the denominator is negative, so it remain to show that the numerator is positive. Using the expressions from Appendix A.2.1 and collecting terms, one can rewrite the numerator as follows:

$$Num = \sum_{i=5,7,9\dots}^{\infty} \Bigg[\overbrace{\frac{64 \cdot 3^i - 192}{i!} - \frac{3 \cdot 2^{i-2}}{(i-4)!}}^{q_1(i)} + \underbrace{\frac{11 \cdot 2^{i-1} + 2 \cdot 3^{i-2} - 6}{(i-3)!} - \frac{48 \cdot 2^i + 16 \cdot 3^i - 48}{(i-1)!} - \frac{2 \cdot 3^{i-1} - 8 \cdot 2^i + 46}{(i-2)!}}_{q_2(i)} \Bigg] x^i$$

It is easy to verify that $q_1(i) + q_2(i) = 0$ for $i \in \{5, 7\}$ and $q_1(i) + q_2(i) > 0$ for $i \in \{9, 11, 13, 15, 17\}$. I will now show by induction that both $q_1(i)$ and $q_2(i)$ are strictly positive for all $i \geq 19$.

- Showing that $q_1(i)$ is positive is equivalent to showing that $3^{i-1} - 1 > 2^{i-8} i(i-1)(i-2)(i-3)$, but it suffices to show that $3^{i-1} - 1 > 2^{i-8} i^4$. The latter is true for $i = 19$. Assume now that it is true for some i and consider $i+1$: $2^{i-7}(i+1)^4 = 2^{i-8} i^4 \frac{2(i+1)^4}{i^4} < [3^{i-1} - 1] \frac{2(i+1)^4}{i^4} = 3^{i-1} \frac{2(i+1)^4}{i^4} - \frac{2(i+1)^4}{i^4} < 3^i - 1$, which concludes the proof by induction. The first inequality follows from the induction assumption and the second one is due to $\frac{2(i+1)^4}{i^4} < 3$ for $i \geq 17$.

- Showing that $q_2(i)$ is positive is equivalent to showing that $(11 \cdot 2^{i-1} + 2 \cdot 3^{i-2} - 6)(i-1)(i-2) > 48 \cdot 2^i + 16 \cdot 3^i - 48 + (2 \cdot 3^{i-1} - 8 \cdot 2^i + 46)(i-1)$, but it suffices to show that $(2 \cdot 3^{i-2} - 6)(i-1)(i-2) > 48 \cdot 2^i + 16 \cdot 3^i + 2 \cdot 3^{i-1}(i-1) + 46(i-1)$. It can be easily shown by induction that each summand on the RHS of the inequality is strictly smaller than $1/4$ of the term on the LHS for all $i \geq 19$, so that the inequality holds true for $i \geq 19$.

A.2.11 Lemma 4

Let $\rho, \kappa, \sigma, \lambda > 0$. Then $\frac{\partial \Theta(2\overline{z}_0,\mu)}{\partial \mu}\Big|_{\mu=0} > 0$ and $\frac{\partial M(2\overline{z}_0,\mu)}{\partial \mu}\Big|_{\mu=0} < 0$.

Proof. Note that expressions for $\frac{\partial \Theta(2\overline{z}_0,\mu)}{\partial \mu}\Big|_{\mu=0}$ and $\frac{\partial M(2\overline{z}_0,\mu)}{\partial \mu}\Big|_{\mu=0}$ can be split in several parts:

$$\frac{\partial \Theta(2\overline{z}_0,\mu)}{\partial \mu}\bigg|_{\mu=0} = \frac{1}{2(q(\overline{z}_0)-2)}\left[\frac{2(\overbrace{q(\overline{z}_0)-2-\alpha^2\overline{z}_0^2}^{A})}{\lambda} + \frac{4(\overbrace{\alpha\overline{z}_0 q(\overline{z}_0)-p(\overline{z}_0)}^{B})}{p(\overline{z}_0)}\left(\frac{\partial \hat{z}(0)}{\partial \mu} - \frac{\partial \underline{z}(0)}{\partial \mu}\right)\right]$$

$$\frac{\partial M(2\overline{z}_0,\mu)}{\partial \mu}\bigg|_{\mu=0} = -\frac{\overbrace{2q(2\overline{z}_0) - 8q(\overline{z}_0) + 12 - \alpha^3\overline{z}_0^3 p(\overline{z}_0)}^{C}}{2\lambda^2(q(\overline{z}_0)-2)^2}$$

$$-\alpha\overline{z}_0\frac{\overbrace{\alpha\overline{z}_0 q(3\overline{z}_0)-2p(3\overline{z}_0)+2p(2\overline{z}_0)+3\alpha\overline{z}_0 q(\overline{z}_0)+2p(\overline{z}_0)-8\alpha\overline{z}_0}^{D}}{2\lambda p(\overline{z}_0)^2(q(\overline{z}_0)-2)^2}\left(\frac{\partial \hat{z}(0)}{\partial \mu} - \frac{\partial \underline{z}(0)}{\partial \mu}\right)$$

So it remains to show that $A, B, C, D > 0$. Using expressions from Appendix A.1.3, letting $x := \alpha\overline{z}(0)$ and collecting terms yields:

$$A = 8\sum_{4,6,8\ldots}^{\infty}\frac{x^i}{i!} > 0$$

$$B = 2\sum_{3,5,7\ldots}^{\infty}\frac{(i-1)x^i}{i!} > 0$$

$$C = \sum_{4,6,8\ldots}^{\infty}[\underbrace{4\cdot 2^i - 16 - 2i(i-1)(i-2)}_{f_C(i)}]\frac{x^i}{i!}$$

$$D = \sum_{3,5,7\ldots}^{\infty}[\underbrace{2i\cdot 3^{i-1} - 4\cdot 3^i + 4\cdot 2^i + 6i + 4}_{f_D(i)}]\frac{x^i}{i!}$$

It remains to show that $f_C(i)$ and $f_D(i)$ are non-negative and take strictly positive values for some i.

- $f_C(i)$. Note that $f_C(4) = f_C(6) = 0$ and $f_C(8) > 0$. It suffices to show that $4 \times 2^i > 16 + 2i^3$ for all $i \geq 10$. Note first that it is true for $i = 10$. Second, assume by induction that it holds for some i and consider $i+1$:

$4 \times 2^{i+1} = 8 \times 2^i > 32 + 4i^3 > 16 + 2(i+1)^3$, where the first inequality is due to the induction assumption and the second one is true since $4i^3 > +2(i+1)^3$ for $i \geq 4$.

- $f_D(i)$. Note that $f_C(3) = f_C(5) = 0$. It suffices to show that $2i \times 3^{i-1} - 4 \times 3^i > 0$ for all $i \geq 7$. This holds since $2i \times 3^{i-1} - 4 \times 3^i = 3^{i-1}(2i - 12) > 0$ for all $i > 6$.

A.2.12 Proofs of several results regarding $\Theta(\delta, \mu)$ and $M(\delta, \mu)$

- **Result 1**

$$\lim_{\delta \to 0} \frac{\partial A_\Theta(\delta, 0)}{\partial \mu} = \frac{2\overline{z}_0}{\sigma^2}$$

Recall that for a small shock ($\delta < \overline{z}_0$):

$$\Theta(\delta, 0) = \frac{1}{6\overline{z}_0^2} \delta^2 (\delta + 3\overline{z}_0)$$

$$\frac{\partial \Theta(\delta, 0)}{\partial \mu} = \frac{\delta^2 (6\overline{z}_0^2 - \delta^2 - 2\delta \overline{z}_0)}{12\sigma^2 \overline{z}_0^2} - \frac{\delta^3}{6\overline{z}_0^3} \frac{\partial \Delta^+(0)}{\partial \mu}$$

So that:

$$\frac{\partial A_\Theta(\delta, 0)}{\partial \mu} = \frac{2}{\Theta(\delta, 0)} \frac{\partial \Theta(\delta, 0)}{\partial \mu} = 2 \left[\frac{(6\overline{z}_0^2 - \delta^2 - 2\delta \overline{z}_0)}{2\sigma^2 (\delta + 3\overline{z}_0)} - \frac{\delta}{\overline{z}_0 (\delta + 3\overline{z}_0)} \frac{\partial \Delta^+(0)}{\partial \mu} \right]$$

Taking the limit as $\delta \to 0$ provides the result.

- **Result 2**

$$\Theta(\delta, \mu) \approx \begin{cases} (1 + \frac{\overline{z}_0}{\sigma^2} \mu) \Theta(\delta, 0) & \text{for } \delta > 0 \\ (1 - \frac{\overline{z}_0}{\sigma^2} \mu) \Theta(\delta, 0) & \text{for } \delta < 0 \end{cases}$$

First order approximation of $\Theta(\delta, \mu)$ with respect to drift μ is given by:

$$\Theta(\delta, \mu) \approx \Theta(\delta, 0) + \frac{\partial \Theta(\delta, 0)}{\partial \mu} \mu$$

Now approximate each term to second order with respect to positive shock $\delta > 0$:

$$\Theta(\delta, 0) \approx \frac{\delta^2}{2\overline{z}_0}$$

$$\frac{\partial \Theta(\delta, 0)}{\partial \mu} \approx \frac{\delta^2}{2\sigma^2}$$

Then:

$$\Theta(\delta,\mu) \approx \frac{\delta^2}{2\overline{z}_0} + \frac{\delta^2}{2\sigma^2}\mu = \left(1 + \frac{\overline{z}_0}{\sigma^2}\mu\right)\frac{\delta^2}{2\overline{z}_0} \approx \left(1 + \frac{\overline{z}_0}{\sigma^2}\mu\right)\Theta(\delta,0)$$

The result for $\delta < 0$ can be shown analogously, with the only difference that second order approximation of $\Theta(\delta,0)$ is given by: $\Theta(\delta,0) \approx -\frac{\delta^2}{2\overline{z}_0}$.

- **Result 3**

$$\lim_{\delta\to 0} \frac{\partial A_\Theta(\delta,0)}{\partial\mu} > \frac{\partial A_I(0)}{\partial\mu}$$

Using expressions for asymmetries, above relation is equivalent to:

$$\frac{2\overline{z}_0}{\sigma^2} > \frac{2}{\overline{z}_0}\frac{\partial\Delta^+(0)}{\partial\mu} \iff \frac{\overline{z}_0^2}{\sigma^2} > \frac{\partial\Delta^+(0)}{\partial\mu}$$

which follows from Lemma 1.

- **Result 4**

$$M(\delta,\mu) \approx \begin{cases} (1 - \frac{|\delta|}{\sigma^2}\mu)M(\delta,0) & \text{for } \delta > 0 \\ (1 + \frac{|\delta|}{\sigma^2}\mu)M(\delta,0) & \text{for } \delta < 0 \end{cases}$$

First order approximation of $M(\delta,\mu)$ with respect to drift μ is given by:

$$M(\delta,\mu) \approx M(\delta,0) + \frac{\partial M(\delta,0)}{\partial\mu}\mu$$

Now approximate each term to second order with respect to shock δ:

$$M(\delta,0) \approx \frac{\overline{z}_0^2\delta}{6\sigma^2}$$
$$\frac{\partial M(\delta,0)}{\partial\mu} \approx -\frac{\overline{z}_0^2\delta^2}{6\sigma^4}$$

Then for $\delta > 0$:

$$M(\delta,\mu) \approx \frac{\overline{z}_0^2\delta}{6\sigma^2} - \frac{\overline{z}_0^2\delta^2}{6\sigma^4}\mu = \left(1 - \frac{|\delta|}{\sigma^2}\mu\right)\frac{\overline{z}_0^2\delta}{6\sigma^2} \approx \left(1 - \frac{|\delta|}{\sigma^2}\mu\right)M(\delta,0)$$

The result for $\delta < 0$ is analogous and immediate.

A.3 Empirics

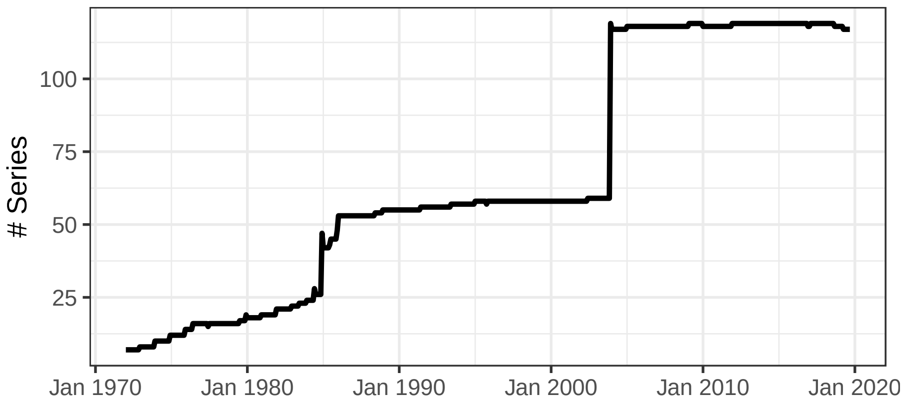
A.3.1 Sectoral Data Construction and Description

I use monthly sectoral data on industrial production index (IP) provided by the Board of Governors of the Federal Reserve System (Industrial Production and Capacity Utilization - G.17). The original data set spans between January 1972 and October 2019 and contains 224 sectors at different levels of aggregation, corresponding to 3-, 4-, 5-, and 6-digit NAICS sectors, and some series contain several NAICS categories. The data on Producer Price Index (PPI) is taken from the Bureau of Labor Statistics, where each series corresponds to a certain NAICS sector, but time spans vary greatly across sectors.

I pair the two data sets in the following way. First, I only keep IP series at the most disaggregated NAICS level (by e.g. omitting 3-digit sectors if a 5-digit sector within that 3-digit sector is present in the data). I also remove series containing several sectors if data in each of these sectors is available individually. This reduces the IP data set to 119 series. Second, for each series in the IP data I produce a corresponding PPI series. If an IP series contains only one NAICS sector, the pairing is straightforward. If an IP series contains several NAICS sectors, I compute a simple average of PPI in these sectors.

An issue with the resulting data set is the sparsity of PPI data. Figure A.3.1 plots the number of series with non-missing values for PPI over time: Since I am interested in estimating impulse responses to identified monetary

Figure A.3.1: Availability of PPI data

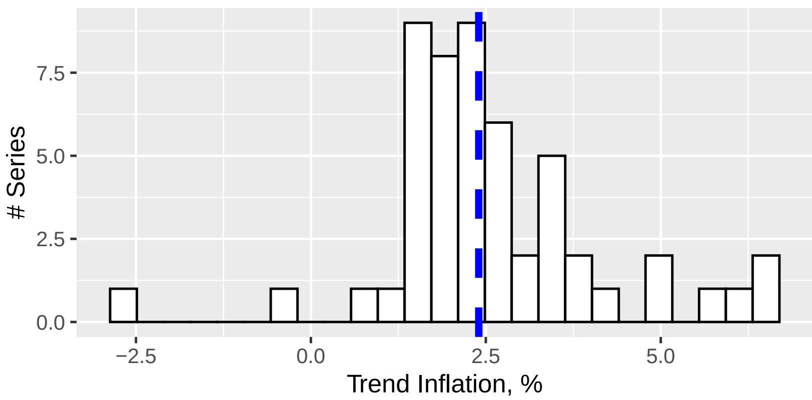


Number of series with non-missing values for PPI over time.

shocks, it is crucial to have a balanced panel to ensure that responses of each series are estimated on the same sample of shocks. Two dates stand out as potential candidates for truncation: January 1986 and January 2004. The latter provides a panel that is approximately twice as large and twice as short as the former one. I restrict the sample to series starting in January 1986, as estimating impulse responses on very short series may be problematic. I also omit one series that has a prolonged period of missing values. The resulting sample contains 52 series, covering the manufacturing sector (NAICS sectors in 31 - 33), logging (NAICS 1133), mining, quarrying, and oil and gas extraction (NAICS sectors in 21), and newspaper, periodical, book, and directory publishers (NAICS sectors in 5111). I set the end date to December 2017, as several series have missing values in 2018 and later. Note that this provides the starting sectoral dataset, which is then further truncated depending on availability of aggregate variables and identified monetary shocks (e.g. in the baseline estimation I consider the period between February 1990 and January 2013).

For each series I compute trend inflation as the average annual PPI growth rate over the entire period. Figure A.3.2 shows the cross-sectional distribution of the estimates. The blue dashed line depicts the median, which is used

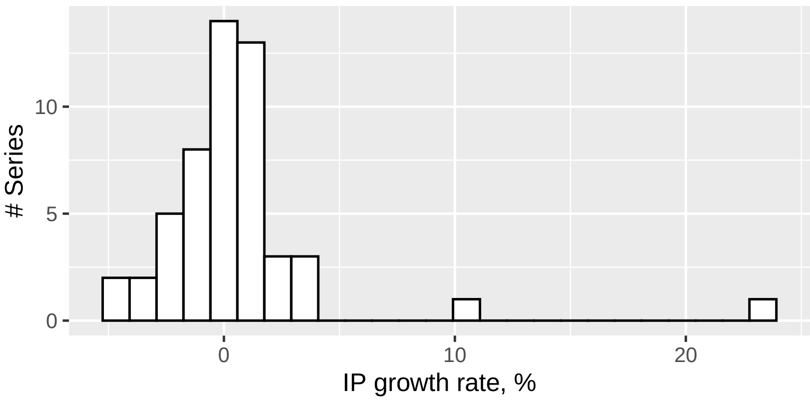
Figure A.3.2: Distribution of Trend Inflation



The blue dashed line shows the median.

to separate series into 'high' and 'low' trend inflation groups. Figure A.3.3 shows the cross-sectional distribution of average annual IP growth rates. The two sectors with the highest production growth are communications equipment manufacturing (3342) and semiconductor and other electronic component manufacturing (3344). These are also the two sectors with negative trend inflation in Figure A.3.2, and are excluded in the baseline estimation

Figure A.3.3: Distribution of IP Growth Rates



by trimming the top and bottom 2.5% of the distribution of trend inflation. The two sectors with the highest trend inflation are drilling oil and gas wells (213111) and petroleum refineries (32411). The two sectors with the largest negative IP growth are leather and hide tanning and finishing (316) and newspaper publishers (51111).

A.3.2 Billion Prices Project Data

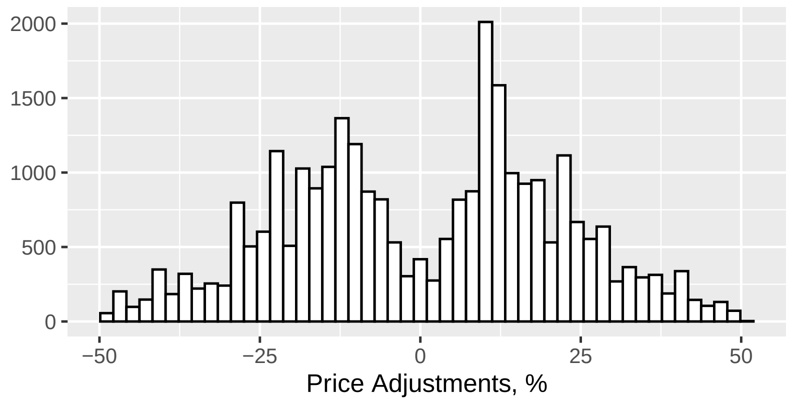
Table A.1: Summary Statistics, Billion Prices Project Data

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Asym. $\log \frac{\Delta^{+} p_i}{\Delta^{-} p_i}$	1,924	−0.019	0.258	−0.852	−0.184	0.136	1.164
Drift μ	1,924	0.002	0.008	−0.022	−0.002	0.007	0.027
Drift μ (alt.)	1,924	0.003	0.008	−0.022	−0.003	0.008	0.028
σ^2	1,924	0.027	0.017	0.001	0.015	0.035	0.106
Frequency	1,924	0.597	0.193	0.369	0.461	0.681	1.402
Std. Dev.	1,924	0.211	0.055	0.043	0.176	0.248	0.375

Each statistic is calculated at the item level. Asymmetry is measured as the log-ratio between magnitudes of positive and negative price adjustments. Std. Dev. stands for the standard deviation of price adjustments. Both measures of drift, idiosyncratic volatility σ^2, and the frequency of price adjustments are computed at a monthly rate.

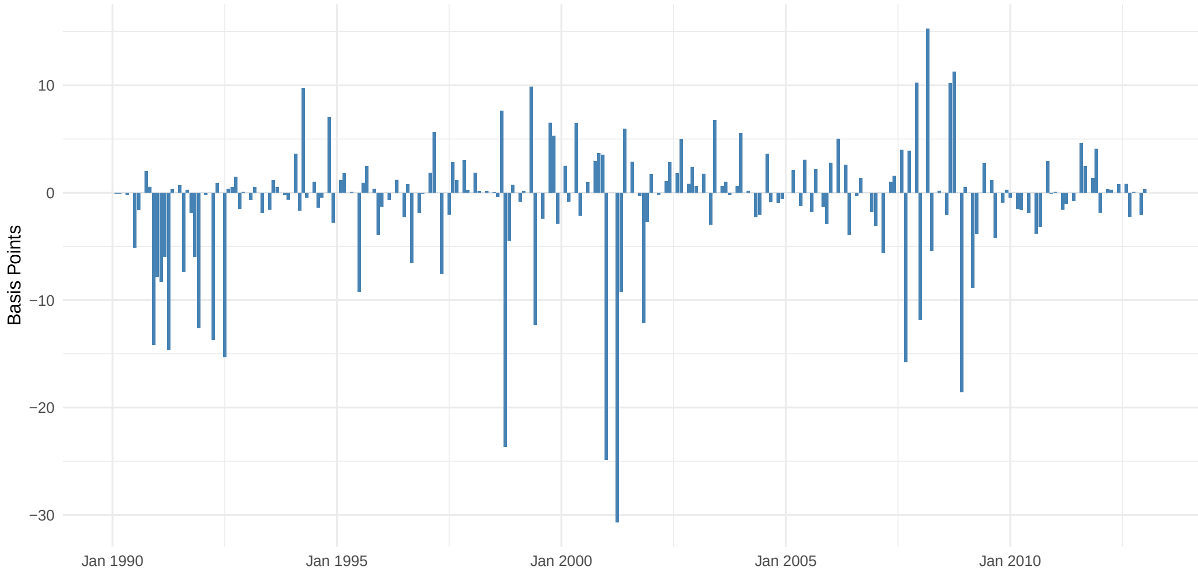
Figure A.3.4 shows the distribution of price adjustments in the sample. The distribution speaks in favor of fixed costs of price adjustment, since small price adjustments are less frequent than adjustments of moderate size. Typically, such pattern can not be observed in lower frequency data, such as biweekly or monthly, which highlights the importance of using daily data when working with price adjustments.

Figure A.3.4: Distribution of Price Adjustments



A.3.3 Monetary Policy Shocks

Figure A.3.5: Monetary Policy Shocks



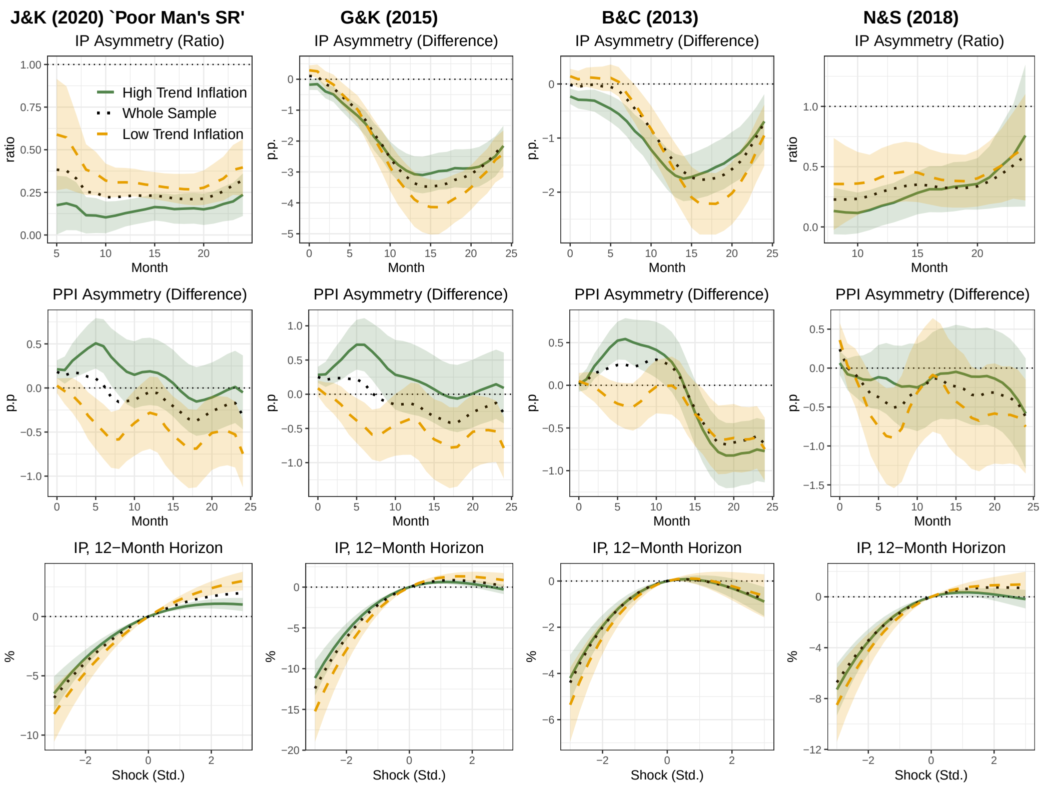
Monetary policy shocks, as estimated by Jarociński and Karadi (2020)

A.3.4 Robustness Checks

Alternative Shock Measures

To ease comparison, I only plot asymmetries (1st row: industrial production, 2nd row: PPI) and non-linear IP responses at a 12-months horizon (3rd row) for each alternative shock measure by column: (1) Jarociński and Karadi (2020) 'poor man's sign restrictions', (2) Gertler and Karadi (2015), (3) Barakchian and Crowe (2013), (4) Nakamura and Steinsson (2018). Whenever possible, I use the preferred measure of asymmetry (ratio), otherwise I compute asymmetry as a difference. The main results of the paper remain generally valid. Asymmetry in the IP responses relates negatively to trend inflation, although results are weaker when measuring asymmetry as a difference. Asymmetry in the PPI responses relates positively to trend inflation. Large positive shocks tend to cause contractions in IP when trend inflation is high, but not so much when trend inflation is low.

Figure A.3.6: Main Results under Alternative Shock Series



Measurement Error

Figure A.3.7: Asymmetry for a Top-30% / Bottom-30% Split

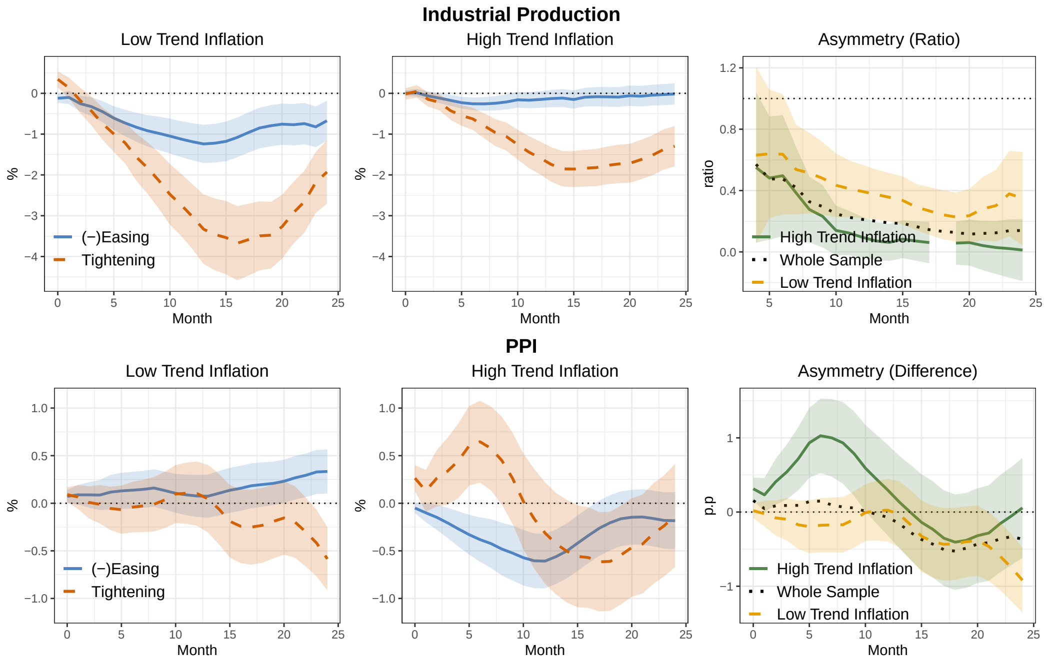
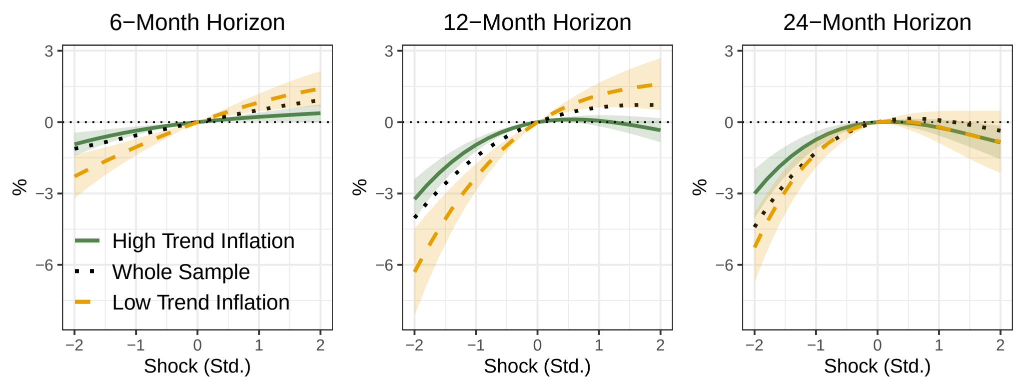


Figure A.3.8: Non-Linearity of Industrial Production Responses for a Top-30% / Bottom-30% Split



Excluding Great Recession and ZLB period

Figure A.3.9: Asymmetry, Sample Until June 2008

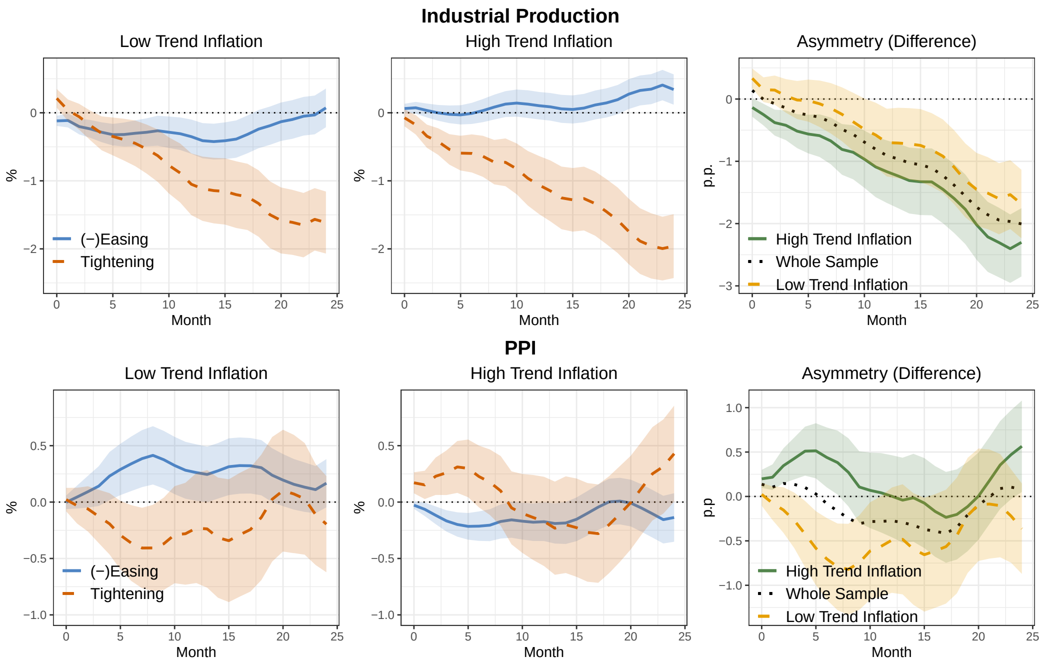
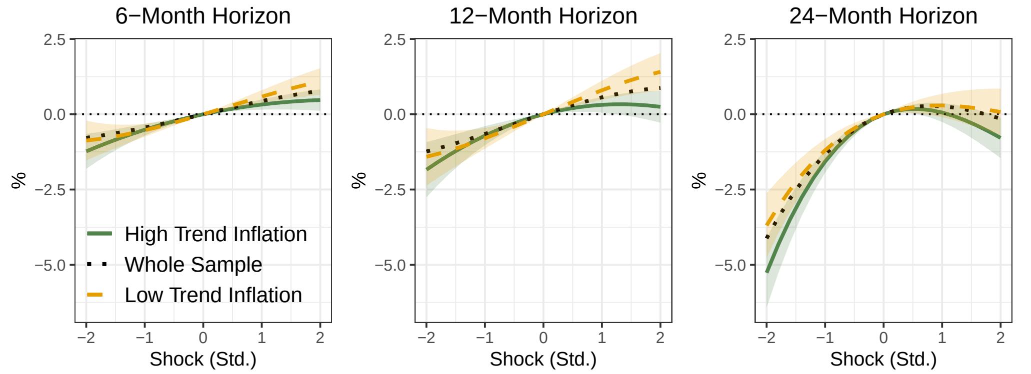


Figure A.3.10: Non-Linearity of Industrial Production Responses, Sample Until June 2008



Alternative Trimming

Figure A.3.11: Asymmetry, Trimming Top and Bottom 15%

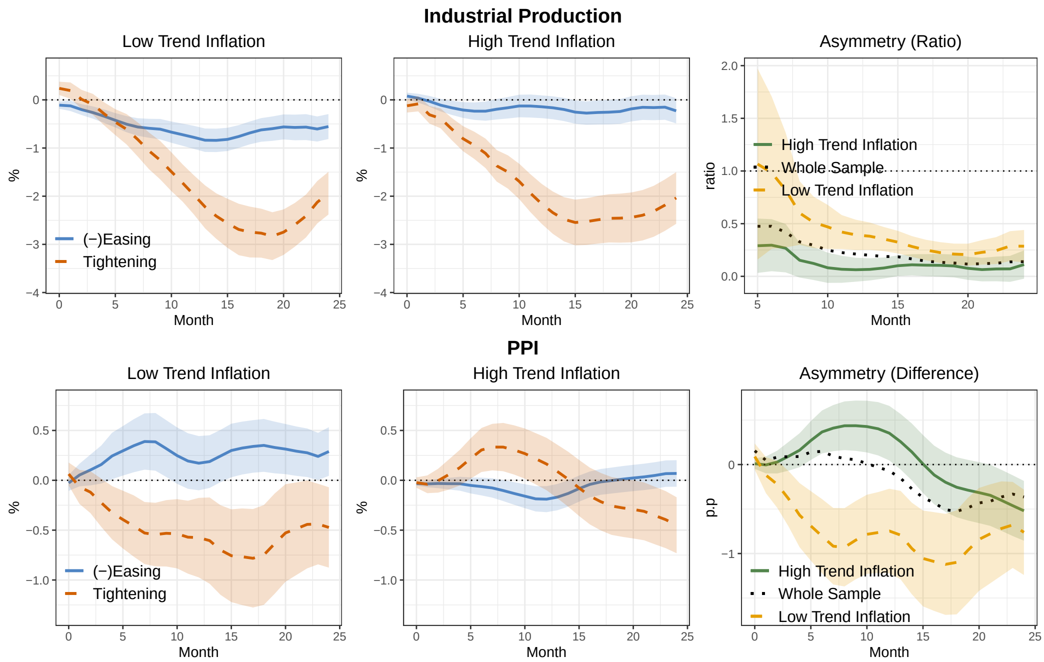


Figure A.3.12: Non-Linearity of Industrial Production Responses, Trimming Top and Bottom 15%

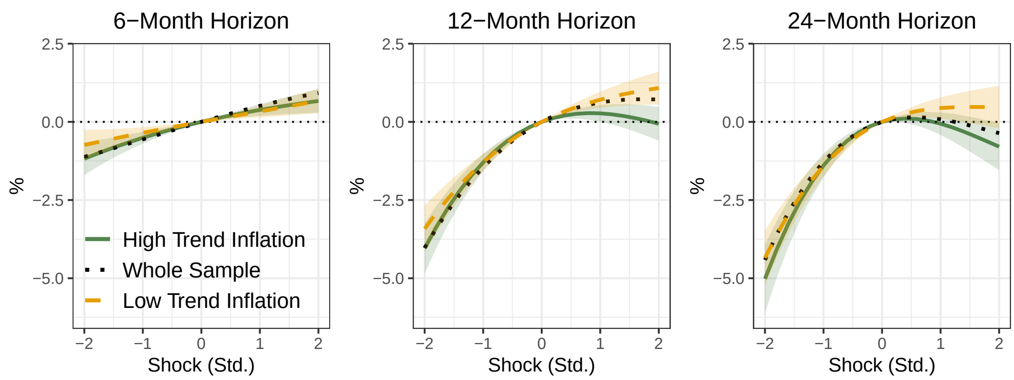


Figure A.3.13: Asymmetry, No Trimming

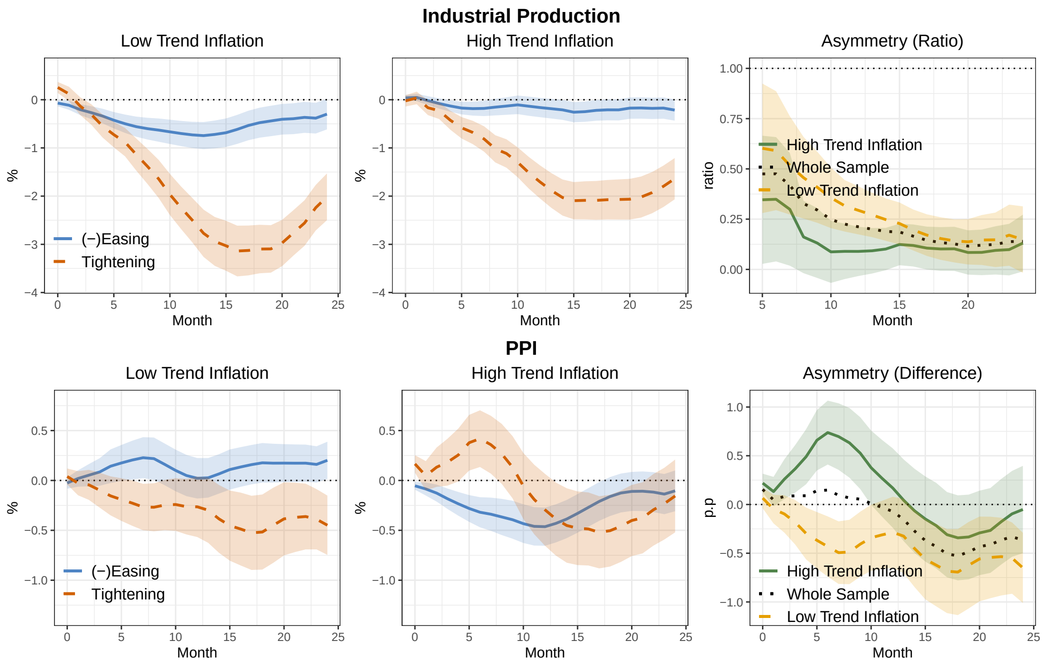
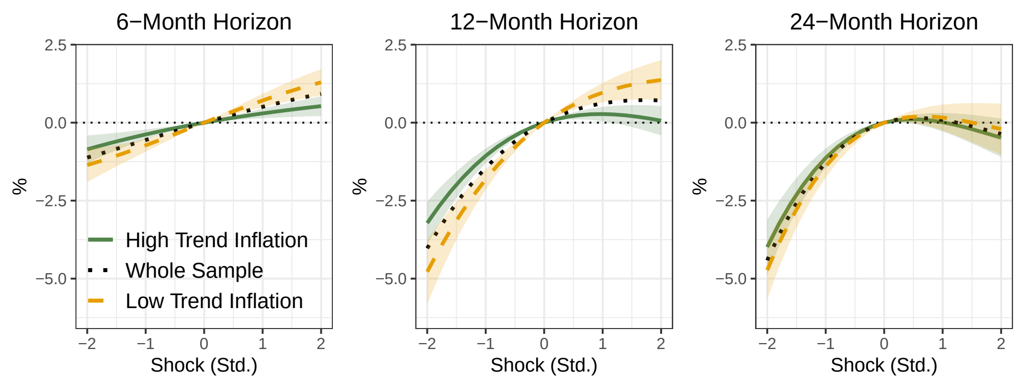
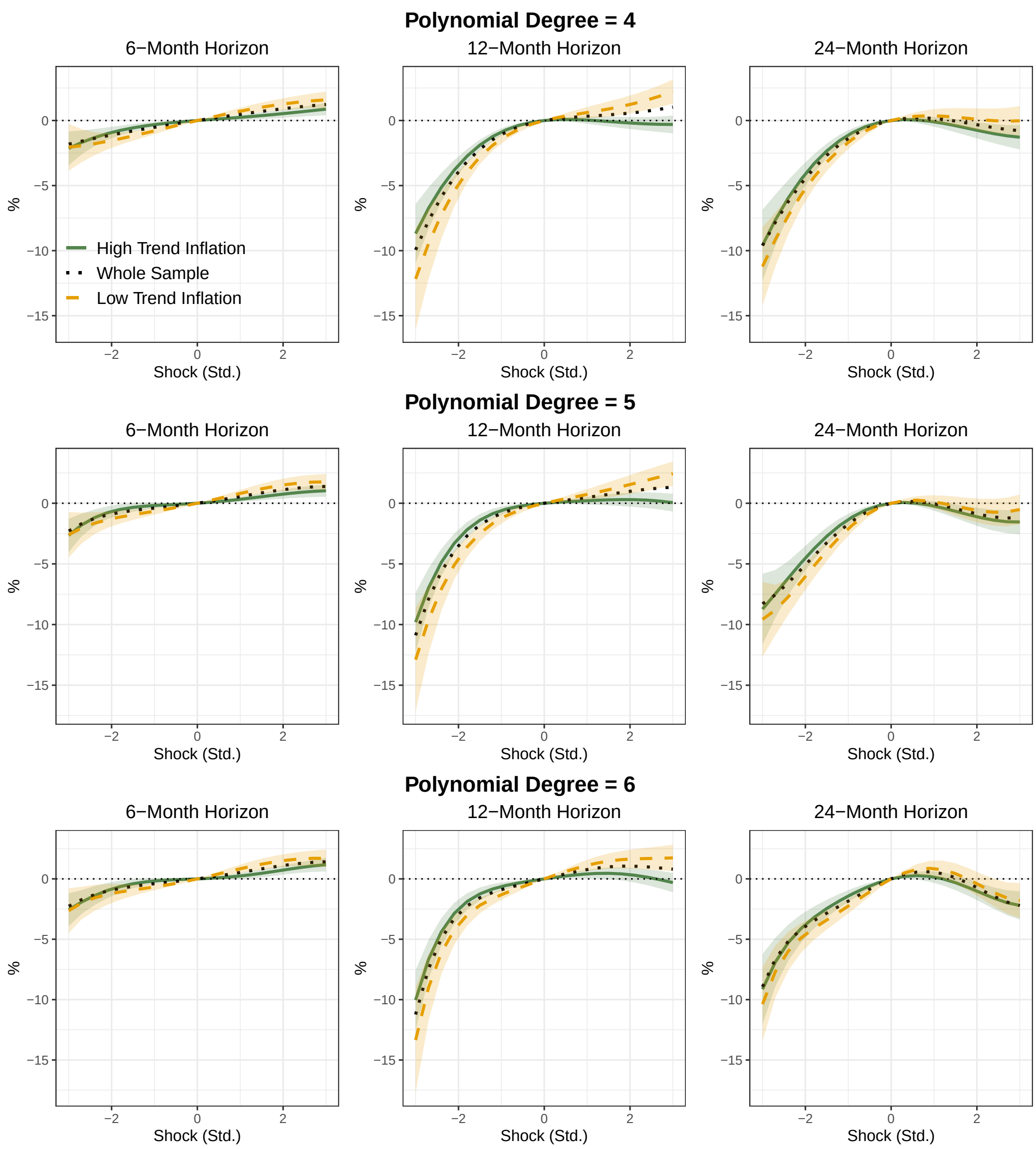


Figure A.3.14: Non-Linearity of Industrial Production Responses, No Trimming



Varying Polynomial Degree

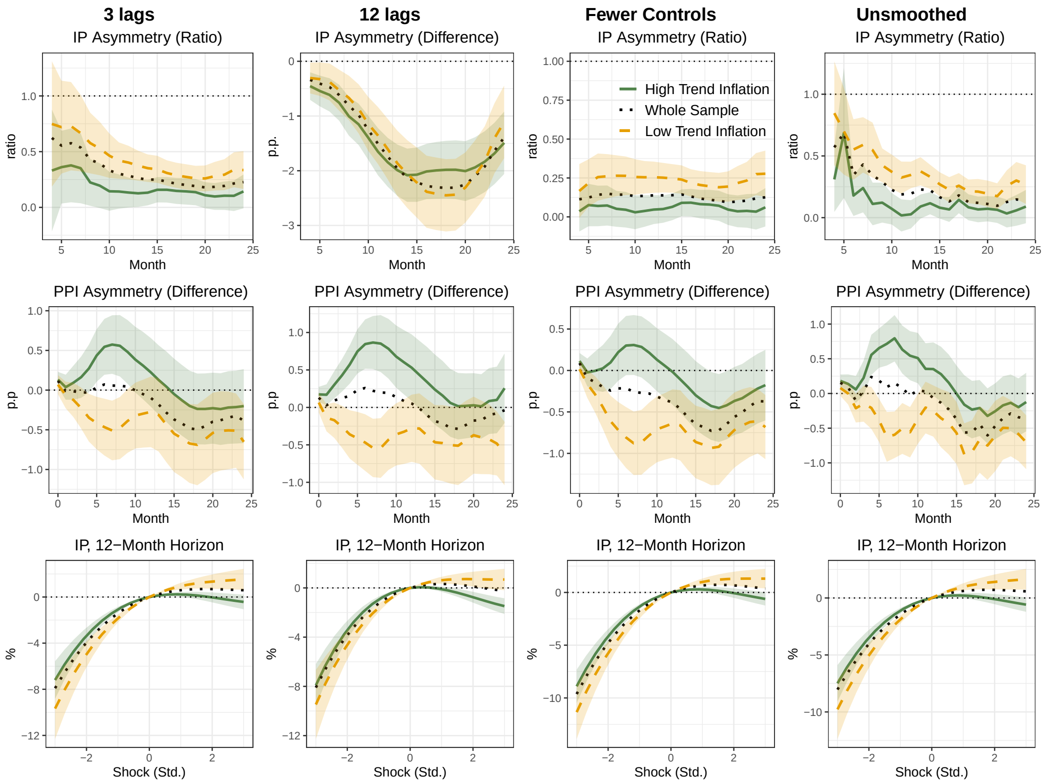
Figure A.3.15: Non-Linearity of Industrial Production Impulse Responses



Other

To ease comparison, I only plot asymmetries (1st row: industrial production, 2nd row: PPI) and non-linear IP responses at a 12-months horizon (3rd row) for each alternative specification by column: (1) number of lags is set to 3, (2) number of lags is set to 12, (3) set of controls consists of a time trend and lags of the dependent variable, monetary shock and effective federal funds rate, (4) unsmoothed impulse responses. Whenever possible, I use the preferred measure of asymmetry (ratio), otherwise I compute asymmetry as difference. The main results of the paper are unchanged.

Figure A.3.16: Other Robustness Checks



A.4 General Equilibrium

A.4.1 Equilibrium along the transition path

Deterministic dynamics after transitory shocks, considered in this paper, introduce three changes relative to the stationary equilibrium. First, the changing markup makes firms profits time dependent. Second, the drift in firms optimal price is also affected by the moving markups and the nominal wage. Third, aggregate consumption can not be omitted from the firms problem, as it is no longer constant.

Denote the time-dependent drift in firms optimal price by $\mu_t = \frac{d\log M_t + d\log\left(\frac{\theta_t}{\theta_t - 1}\right)}{dt}$. The value function of a firm becomes time-dependent:

$$(\rho+\lambda)v(z,t) = \pi(z,t) + \lambda v(\hat{z},t) - \mu_t v_z(z,t) + \frac{1}{2}\sigma^2 v_{zz}(z,t) + v_t(z,t)$$

as well as the distribution of price gaps:

$$f_t(z,t) = \mu_t f_z(z,t) + \frac{1}{2}\sigma^2 f_{zz}(z,t) - \lambda f(z,t)$$

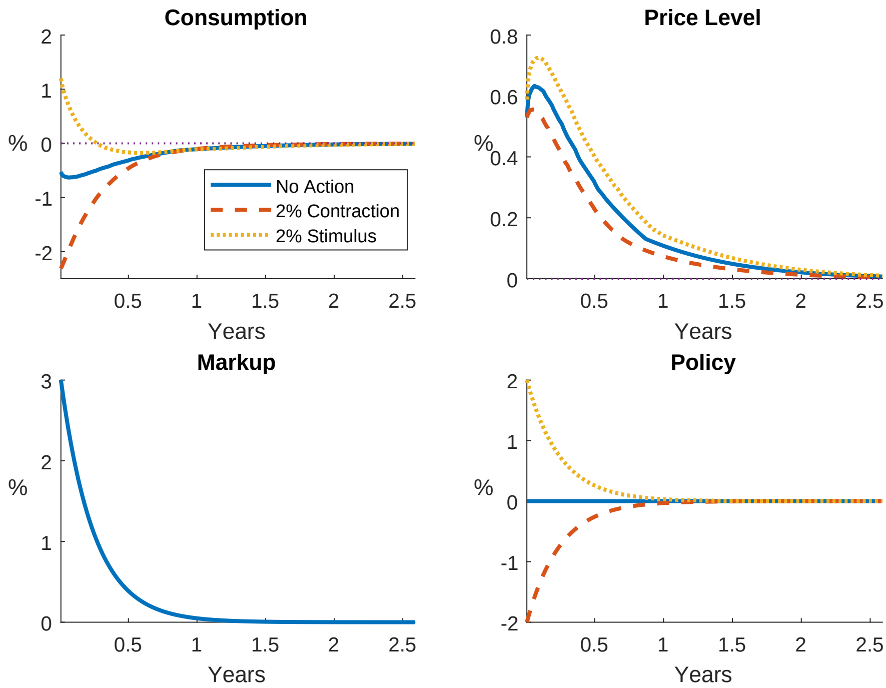
Time-dependent profit and cost functions are:

$$\pi(z,t) = \left(\frac{\alpha\theta_t C_t}{\theta_t - 1}\right)^{1-\theta_t} e^{-\theta_t z}\left(e^z - \frac{\theta_t - 1}{\theta_t}\right)$$
$$c(z,t) = \kappa\left(\frac{\alpha\theta_t C_t}{\theta_t - 1}\right)^{1-\theta_t} e^{(1-\theta_t)z}$$

All other equilibrium objects can be computed as before, substituting constant variables with time-dependent ones. When solving for the transition dynamics, I follow the numerical approach of Achdou et al. (2017).

A.4.2 Dynamics after Policy Interventions

Figure A.4.1: Markup Shock and Policy Response



Impulse responses of consumption, price level and markup to a 3% markup shock and policy intervention. Solid blue lines correspond to a zero monetary response, dashed red lines – to a 2% contraction, dotted yellow lines – to a 2% expansion. Consumption and markup responses are in terms of percent deviations from the steady state, price level responses are in terms of percent deviations from the trend.

A.4.3 Alternative Calibration

I now target the same values of price adjustment frequency and average size of adjustment, but consider a lower target for kurtosis, setting it to 3. The calibrated values in annual terms for σ, κ and λ are now 0.142, 0.048 and 1.03. The next two figures show policymaker's frontiers after 3% and 10% markup shocks, same as those considered under the baseline calibration.

Figure A.4.2: Frontiers, Small Shock

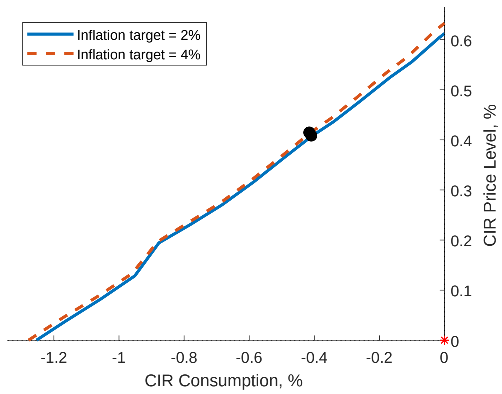
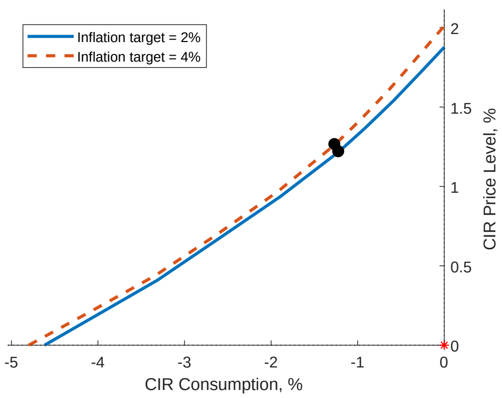


Figure A.4.3: Frontiers, Large Shock



Increasing inflation target from 2% to 4% amplifies the response to the markup shock by 1.2% when the shocks is small (2%) and by 4.1% when the shocks is large (10%). At the same time, the curvature of the frontier increases by 7.4% for the small shock and by 10.2% for the large shock. Thus, the effect of trend inflation remains quantitatively sizable under an alternative calibration.

A.4.4 Using Inflation CIR

Here I consider an alternative frontier of the monetary authority, defined in terms of the usual consumption CIR and a cumulative response of absolute values of inflation: $\int_0^\infty |\pi_t - \mu| dt$. The next two figures show policymaker's frontiers after 3% and 10% initial markup disturbances, same as those considered under the baseline specification.

Figure A.4.4: Frontiers, Small Shock

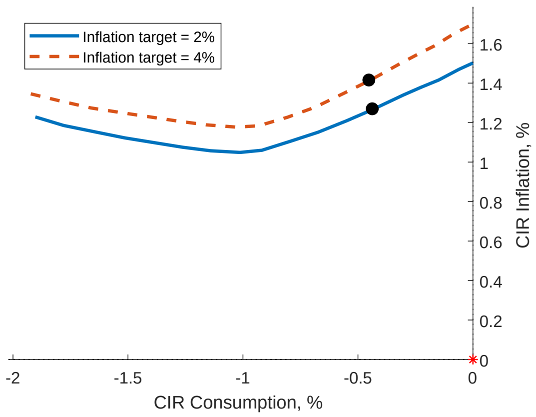
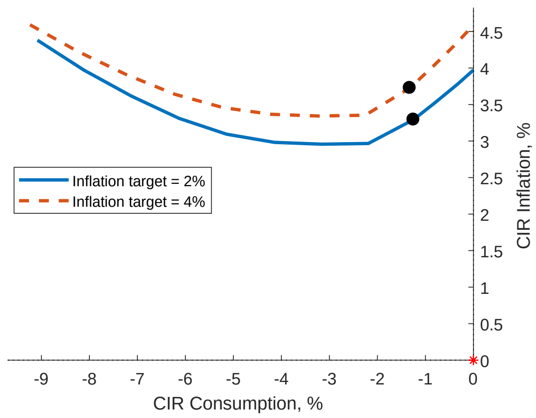


Figure A.4.5: Frontiers, Large Shock



In both cases a higher trend inflation leads to a larger initial response to the markup shock in terms of both consumption and inflation deviations. In addition, monetary authority becomes more constrained in stabilizing inflation, as some levels of inflation CIR become infeasible (the red dashed lines lie above the blue ones).

Appendix B

Chapter 2

B.1 Key Expressions

I derive some key expressions, which are used in the main text and in the proofs in Appendix B.2.

Integrating by parts, the return on purchasing contract φ can be rewritten as:

$$\begin{aligned}
R_c(i,\varphi) &= \frac{\mathbb{E}_i[\min(y,\varphi)]}{q(\varphi)} = \frac{\int_{\underline{c}}^{\varphi} y dF_i(y) + \varphi(1-F_i(\varphi))}{q(\varphi)} \\
&= \frac{yF_i(y)\big|_{\underline{c}}^{\varphi} - \int_{\underline{c}}^{\varphi} F_i(y)dy + \varphi(1-F_i(\varphi))}{q(\varphi)} = \frac{\varphi - \int_{\underline{c}}^{\varphi} F_i(y)dy}{q(\varphi)} \\
&= \frac{\int_{\underline{c}}^{\varphi}(1-F_i(y))dy + \underline{c}}{q(\varphi)}
\end{aligned}$$

Rewriting in the same way the return on buying the asset on margin with contract φ yields:

$$\begin{aligned}
R_y(i,\varphi) &= \frac{\mathbb{E}_i[\max(y-\varphi,0)]}{p-q(\varphi)} = \frac{\int_{\varphi}^{\bar{c}} y dF_i(y) - \varphi(1-F_i(\varphi))}{p-q(\varphi)} \\
&= \frac{yF_i(y)\big|_{\varphi}^{\bar{c}} - \int_{\varphi}^{\bar{c}} F_i(y)dy - \varphi(1-F_i(\varphi))}{p-q(\varphi)} = \frac{(\bar{c}-\varphi) - \int_{\varphi}^{\bar{c}} F_i(y)dy}{p-q(\varphi)} \\
&= \frac{\int_{\varphi}^{\bar{c}}(1-F_i(y))dy}{p-q(\varphi)}
\end{aligned}$$

First order conditions are then:

$$\begin{aligned}
\frac{\partial R_c(i,\varphi)}{\partial \varphi} &= \frac{(1-F_i(\varphi))q(\varphi) - q'(\varphi)R_c(i,\varphi)q(\varphi)}{q(\varphi)^2} = 0 \iff \\
\frac{1-F_i(\varphi)}{R_c(i,\varphi)} &= q'(\varphi)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial R_y(i,\varphi)}{\partial \varphi} &= \frac{-(1-F_i(\varphi))(p-q(\varphi)) + q'(\varphi)R_y(i,\varphi)(p-q(\varphi))}{(p-q(\varphi))^2} = 0 \iff \\
\frac{1-F_i(\varphi)}{R_y(i,\varphi)} &= q'(\varphi)
\end{aligned}$$

B.2 Proofs

Lemma 1: Proof.. *Assume that agent i prefers to buy the asset unleveraged over cash and any contract:* $\frac{\mathbb{E}_i[y]}{p} \geq 1$, $\frac{\mathbb{E}_i[y]}{p} \geq R_c(i,\varphi)\ \forall \varphi$. *Rewrite the latter* *equation:*

$$\frac{\int_{\underline{c}}^{\overline{c}} ydF_i(y)}{p} \geq \frac{\int_{\underline{c}}^{\varphi} ydF_i(y) + \varphi(1 - F_i(\varphi))}{q(\varphi)} \quad \forall \varphi$$

It follows that:

$$q(\varphi) \int_{\underline{c}}^{\overline{c}} ydF_i(y) \geq p \left[\int_{\underline{c}}^{\varphi} ydF_i(y) + \varphi(1 - F_i(\varphi)) \right] \; \forall \varphi \Longrightarrow$$

$$p \left[\int_{\underline{c}}^{\overline{c}} ydF_i(y) - \left(\int_{\underline{c}}^{\varphi} ydF_i(y) + \varphi(1 - F_i(\varphi)) \right) \right] \geq (p - q(\varphi)) \int_{\underline{c}}^{\overline{c}} ydF_i(y) \; \forall \varphi \Longrightarrow$$

$$\frac{\int_{\varphi}^{\overline{c}} ydF_i(y) - \varphi(1 - F_i(\varphi))}{p - q(\varphi)} \geq \frac{\int_{\underline{c}}^{\overline{c}} ydF_i(y)}{p} \; \forall \varphi$$

Where the last line follows since $p > q(\varphi)$*. The left-hand side is then* $R_y(i, \varphi)$*. Thus agent* i *weakly prefers to buy the asset on margin.*

Lemma A.

- $R_c(i, \varphi)$ *is strictly increasing in* i *for all* $\varphi > \underline{c}$ *and is constant in* i *for all* $\varphi \leq \underline{c}$
- $R_y(i, \varphi)$ *is strictly increasing in* i *for all* $\varphi < \overline{c}$ *and is zero for all* i *for all* $\varphi \geq \overline{c}$

Proof. *For the first part, consider* $\varphi > \underline{c}$*,* $j > i$*.*

$$R_c(j, \varphi) = \frac{\int_{\underline{c}}^{\varphi} (1 - F_j(y))dy + \underline{c}}{q(\varphi)} > \frac{\int_{\underline{c}}^{\varphi} (1 - F_i(y))dy + \underline{c}}{q(\varphi)} = R_c(i, \varphi)$$

where the inequality is due to first order stochastic dominance, implied by hazard-rate order assumption A1. If $\varphi \leq \underline{c}$*, then* $R_c(i, \varphi) = \frac{\varphi}{q(\varphi)}$ *and thus it is constant in* i*.*

For the second part, similarly, consider $\varphi < \overline{c}$*,* $j > i$*.*

$$R_y(j, \varphi) = \frac{\int_{\varphi}^{\overline{c}} (1 - F_j(y))dy}{p - q(\varphi)} > \frac{\int_{\varphi}^{\overline{c}} (1 - F_i(y))dy}{p - q(\varphi)} = R_y(i, \varphi)$$

where the inequality again follows from assumption A1. The fact that $R_y(i, \varphi) = 0$ *for all* $\varphi \geq \overline{c}$ *is immediate, as in that case* $\max(y - \varphi, 0) = 0$ *for all* $y \in [\underline{c}, \overline{c}]$*.*

Theorem 1: Proof.. *In order to prove Theorem 1, I use the following Lemmas:*

- ***Lemma 1.1***
  If each agent j prefers to buy cash or riskless contract, then so does any agent $i < j$.

- ***Lemma 1.2***
  If agent i prefers to buy the asset leveraged, then any agent $j > i$ strictly[1] prefers to buy the asset leveraged.

- ***Lemma 1.3***
  Risky contracts are traded in any equilibrium

Lemma 1.1 implies that there exists i^* such that any agent $i < i^*$ buys cash or riskless contracts and any agent $j > i^*$ buys risky contracts or the asset leveraged. Lemma 1.2 implies that there exists j^* such that any agent $j > j^*$ buys the asset leveraged and any agent $i < j^*$ buys contracts or cash. Lemma 1.3 then implies that $j^* > i^*$ and any agent i between the two marginal agents buys risky contracts. In the following I provide proofs for the three Lemmas.

Lemma 1.1: Proof.. *First note that riskless contracts provide return of 1 in equilibrium, as they are equivalent to cash. Assume that agent j prefers to buy cash or a riskless contract over the other two options: $R_c(j,\varphi) \leq 1$, $R_y(j,\varphi) \leq 1 \ \forall \varphi > \underline{c}$. From Lemma A it follows immediately that the same inequalities hold for any $i < j$ and thus i also prefers to buy cash or a riskless contract.*

Lemma 1.2: Proof.. *Suppose agent i prefers to buy the asset leveraging with contract φ: $R_y(i,\varphi) \geq R_c(i,\tilde{\varphi}) \ \forall \tilde{\varphi} > \underline{c}$ and $R_y(i,\varphi) \geq 1$. Note that $\varphi < \overline{c}$, as otherwise, by Lemma A, $R_y(i,\varphi) = 0$. Consider agent $j > i$. Let's show that $R_y(j,\varphi) > R_c(j,\tilde{\varphi}) \ \forall \tilde{\varphi} > \underline{c}$, since $R_y(j,\varphi) > 1$ follows immediately from Lemma A.*

$$R_y(j,\varphi) > R_c(j,\tilde{\varphi}) \iff$$

$$\frac{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy}{p-q(\varphi)} > \frac{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy + \underline{c}}{q(\tilde{\varphi})} \iff$$

$$\underbrace{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy}_{A_j} > \underbrace{\frac{p-q(\varphi)}{q(\tilde{\varphi})}}_{\alpha}\underbrace{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy}_{B_j} + \underbrace{\frac{p-q(\varphi)}{q(\tilde{\varphi})}\underline{c}}_{\gamma}$$

In other words, for a fixed $\tilde{\varphi}$, one has to show that $A_j > \alpha B_j + \gamma$, given that $A_i \geq \alpha B_i + \gamma$ and $j > i$. It suffices to show that $\frac{A_j}{B_j} > \frac{A_i}{B_i}$, since then

[1] Strict preference means that j can not be indifferent between purchasing the asset on margin and any other option.

$A_j > \frac{A_i}{B_i}B_j \geq \frac{\alpha B_i+\gamma}{B_i}B_j = \alpha B_j + \gamma\frac{B_j}{B_i} > \alpha B_j + \gamma$, *where the last inequality follows from Lemma A.*

Consider first $\tilde{\varphi} \leq \varphi$.

$$\frac{A_j}{B_j} = \frac{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy} = \frac{\int_{\varphi}^{\overline{c}}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}$$
$$> \frac{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\varphi}^{\overline{c}}(1-F_i(y))dy}{\frac{1-F_j(\tilde{\varphi})}{1-F_i(\tilde{\varphi})}\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} \geq \frac{\int_{\varphi}^{\overline{c}}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} = \frac{A_i}{B_i}$$

where inequalities are due to hazard-rate order assumption A1.

Consider now $\tilde{\varphi} > \varphi$.

$$\frac{A_j}{B_j} = \frac{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy} = \frac{\int_{\tilde{\varphi}}^{\overline{c}}(1-F_j(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy} + \frac{\int_{\varphi}^{\tilde{\varphi}}(1-F_j(y))dy}{\int_{\underline{c}}^{\varphi}(1-F_j(y))dy + \int_{\varphi}^{\tilde{\varphi}}(1-F_j(y))dy}$$
$$= \frac{\int_{\tilde{\varphi}}^{\overline{c}}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy} + \left[1 + \frac{\int_{\underline{c}}^{\varphi}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}{\int_{\varphi}^{\tilde{\varphi}}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}\right]^{-1}$$
$$> \frac{\frac{1-F_j(\tilde{\varphi})}{1-F_i(\tilde{\varphi})}\int_{\tilde{\varphi}}^{\overline{c}}(1-F_i(y))dy}{\frac{1-F_j(\tilde{\varphi})}{1-F_i(\tilde{\varphi})}\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} + \left[1 + \frac{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\underline{c}}^{\varphi}(1-F_i(y))dy}{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\varphi}^{\tilde{\varphi}}(1-F_i(y))dy}\right]^{-1}$$
$$= \frac{\int_{\tilde{\varphi}}^{\overline{c}}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} + \frac{\int_{\varphi}^{\tilde{\varphi}}(1-F_i(y))dy}{\int_{\underline{c}}^{\varphi}(1-F_i(y))dy + \int_{\varphi}^{\tilde{\varphi}}(1-F_i(y))dy} = \frac{\int_{\varphi}^{\overline{c}}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} = \frac{A_i}{B_i}$$

where inequality follows from assumption A1.

Lemma 1.3: Proof.. *First, let's show that in any equilibrium there are contracts traded (potentially riskless). Assume the contrary. That means pessimistic agents buy cash, optimistic agents buy the asset and there is a marginal agent* i^* *who is indifferent:* $\frac{\mathbb{E}_{i^*}[y]}{p} = 1$. *Consider* $j > i^*$ *and some riskless contract* $0 < \varphi \leq \underline{c}$. *In equilibrium* $q(\varphi)$ *is such that* i^* *does not want to buy contract* φ:

$$R_c(i^*, \varphi) \leq 1$$

and j *does not want to sell it:*

$$R_y(j, \varphi) \leq \frac{\mathbb{E}_j[y]}{p}$$

In the following I show that it is impossible to find such $q(\varphi)$. Using the fact that $\varphi \leq \underline{c}$, condition (a) implies:

$$R_c(i^*, \varphi) = \frac{\varphi}{q(\varphi)} \leq 1 \iff \varphi \leq q(\varphi)$$

and condition (b) implies:

$$\begin{aligned} R_y(j, \varphi) = \frac{\mathbb{E}_j[y] - \varphi}{p - q(\varphi)} \leq \frac{\mathbb{E}_j[y]}{p} &\iff \\ p\,(\mathbb{E}_j[y] - \varphi) \leq (p - q(\varphi))\mathbb{E}_j[y] &\iff \\ q(\varphi) \leq \varphi \frac{p}{\mathbb{E}_j[y]} & \end{aligned}$$

Since $j > i^*$, $\mathbb{E}_j[y] > p$ and thus condition (b) requires that $q(\varphi) < \varphi$, which contradicts condition (a). Thus in any equilibrium there are contracts traded.

Assume now that there are no risky contracts traded in equilibrium. That implies agents buy cash with return of 1, lend with a riskless contract (which needs to have the same return as cash) or buy the asset leveraged with a riskless contract. The fact that return on riskless contracts equals 1 implies $q(\varphi) = \varphi \ \forall \varphi \leq \underline{c}$. There is a marginal agent i^* such that for all $i < i^*$, agent i buys cash or riskless contract, and for all $j > i^*$, agent j buys the asset leveraged. Buying the asset and leveraging with riskless contract φ provides expected return:

$$R_y(i, \varphi) = \frac{\mathbb{E}_i[y] - \varphi}{p - \varphi}$$

The marginal buyer is indifferent and thus $p = \mathbb{E}_{i^*}[y]$. Thus for all $j > i^*$, $\mathbb{E}_j[y] > p$ and agents prefer to leverage with the largest riskless contract $\underline{c}$.

Consider now risky contract $\varphi > \underline{c}$ and $j > i^*$. In equilibrium $q(\varphi)$ is such that i^* does not want to buy φ:

$$R_c(i^*, \varphi) \leq 1$$

and j does not want to sell it:

$$R_y(j, \varphi) \leq \frac{\mathbb{E}_j[y] - \underline{c}}{p - \underline{c}}$$

Rewrite condition (b) as:

$$R_y(j,\varphi) \leq \frac{\mathbb{E}_j[y]-\underline{c}}{p-\underline{c}} \iff \frac{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy}{p-q(\varphi)} \leq \frac{\int_{\underline{c}}^{\overline{c}}(1-F_j(y))dy}{p-\underline{c}} \iff$$
$$\frac{p-\underline{c}}{p-q(\varphi)} \leq \frac{\int_{\underline{c}}^{\overline{c}}(1-F_j(y))dy}{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy} \iff \frac{p-\underline{c}}{p-q(\varphi)} \leq 1+\underbrace{\frac{\int_{\underline{c}}^{\varphi}(1-F_j(y))dy}{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy}}_{A_j}$$

Since $p = E_{i^*}[y]$*, condition (a) implies:*

$$R_c(i^*,\varphi) \leq 1 \iff \int_{\underline{c}}^{\varphi}(1-F_i(y))dy+\underline{c} \leq q(\varphi) \implies$$
$$p-q(\varphi) = \int_{\underline{c}}^{\overline{c}}(1-F_i(y))dy+\underline{c}-q(\varphi) \leq \int_{\varphi}^{\overline{c}}(1-F_j(y))dy \iff$$
$$\frac{p-\underline{c}}{p-q(\varphi)} = \frac{\int_{\underline{c}}^{\overline{c}}(1-F_i(y))dy}{\int_{\underline{c}}^{\overline{c}}(1-F_i(y))dy+\underline{c}-q(\varphi)} \geq \frac{\int_{\underline{c}}^{\overline{c}}(1-F_i(y))dy}{\int_{\varphi}^{\overline{c}}(1-F_i(y))dy} = 1+\underbrace{\frac{\int_{\underline{c}}^{\varphi}(1-F_i(y))dy}{\int_{\varphi}^{\overline{c}}(1-F_i(y))dy}}_{A_i}$$

Thus conditions (a) and (b) imply:

$$A_j \geq A_i$$

On the other hand:

$$A_j = \frac{\int_{\underline{c}}^{\varphi}(1-F_j(y))dy}{\int_{\varphi}^{\overline{c}}(1-F_j(y))dy} = \frac{\int_{\underline{c}}^{\varphi}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}{\int_{\varphi}^{\overline{c}}\frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}$$
$$< \frac{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\underline{c}}^{\varphi}(1-F_i(y))dy}{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\varphi}^{\overline{c}}(1-F_i(y))dy} = \frac{\int_{\underline{c}}^{\varphi}(1-F_i(y))dy}{\int_{\varphi}^{\overline{c}}(1-F_i(y))dy} = A_i$$

where the inequality is due to Assumption A1. Thus it is impossible to find $q(\varphi)$ *that satisfies equilibrium conditions and risky contracts are traded in any equilibrium.*

Theorem 2: Proof.. *Assume a riskless contract* $\varphi \leq \underline{c}$ *is traded in equilibrium, i.e. some agent* i *buys the contract and some agent* j *sells it. Assumption A2 ensures contract choices of those buying and selling contracts are given by first order conditions. These imply:*

$$1 = \frac{1-F_i(\varphi)}{R_c(i,\varphi)} = q'(\varphi) = \frac{1-F_j(\varphi)}{R_y(j,\varphi)} = \frac{1}{R_y(j,\varphi)}$$

since $F_i(\varphi) = 0$ for all i and $R_c(i, \varphi) = 1$ as riskless contract is equivalent to cash. As a result, $R_y(j, \varphi) = 1$. Since agent j prefers to buy the asset leveraged, $R_y(j, \varphi) \geq R_c(j, \tilde{\varphi})$ for all $\tilde{\varphi}$. Thus by Lemma A, for any agent $\tilde{j} < j$, $R_c(\tilde{j}, \tilde{\varphi}) < R_c(j, \tilde{\varphi}) \leq 1$ for all $\tilde{\varphi} > \underline{c}$. This means that no agent in the economy is willing to buy risky contracts, since all the agents above j are buying the asset (see the proof of Theorem 1), and all those below prefer cash over risky contracts. This then contradicts Theorem 1 that in any equilibrium there are risky contracts traded.

Lemma B.

- $\frac{R_c(i,\varphi)}{1-F_i(\varphi)}$ *is strictly decreasing in i for all $\varphi \in (\underline{c}, \overline{c})$ and is constant in i for all $\varphi \leq \underline{c}$*
- $\frac{R_y(i,\varphi)}{1-F_i(\varphi)}$ *is strictly increasing in i for all $\varphi < \overline{c}$*

Proof. For the first part, consider $\varphi \in (\underline{c}, \overline{c})$, $j > i$.

$$\begin{aligned}\frac{R_c(j, \varphi)}{1 - F_j(\varphi)} &= \frac{1}{q(\varphi)} \left[\frac{\int_{\underline{c}}^{\varphi} y f_j(y) dy}{1 - F_j(\varphi)} + \varphi \right] < \frac{1}{q(\varphi)} \left[\frac{\int_{\underline{c}}^{\varphi} y f_i(y) \frac{1-F_j(y)}{1-F_i(y)} dy}{1 - F_j(\varphi)} + \varphi \right] \\ &< \frac{1}{q(\varphi)} \left[\frac{\frac{1-F_j(\varphi)}{1-F_i(\varphi)} \int_{\underline{c}}^{\varphi} y f_i(y) dy}{1 - F_j(\varphi)} + \varphi \right] = \frac{R_c(i, \varphi)}{1 - F_i(\varphi)}\end{aligned}$$

where the inequalities follows from hazard-rate order assumption A1. If $\varphi \leq \underline{c}$, then $\frac{R_c(j,\varphi)}{1-F_j(\varphi)} = \frac{\varphi}{q(\varphi)}$ and thus is constant in i.
For the second part, similarly, consider $\varphi < \overline{c}$, $j > i$.

$$\begin{aligned}\frac{R_y(j, \varphi)}{1 - F_j(\varphi)} &= \frac{1}{p - q(\varphi)} \left[\frac{\int_{\varphi}^{\overline{c}} (1 - F_j(y)) dy}{1 - F_j(\varphi)} \right] = \frac{1}{p - q(\varphi)} \left[\frac{\int_{\varphi}^{\overline{c}} \frac{1-F_j(y)}{1-F_i(y)} (1 - F_i(y)) dy}{1 - F_j(\varphi)} \right] \\ &> \frac{1}{p - q(\varphi)} \left[\frac{\frac{1-F_j(\varphi)}{1-F_i(\varphi)} \int_{\varphi}^{\overline{c}} (1 - F_i(y)) dy}{1 - F_j(\varphi)} \right] = \frac{R_y(i, \varphi)}{1 - F_i(\varphi)}\end{aligned}$$

where the inequalities again follow from assumption A1.

Theorem 3: Proof.. *In order to prove Theorem 3, I use the following Lemmas:*

- ***Lemma 3.1***
  If contract φ is traded in equilibrium, then there exists a unique i and a unique j such that agent i buys the contract and agent j sells it.

- ***Lemma 3.2***
  If agent i buys contract φ and agent $j > i$ buys contract $\tilde{\varphi}$, then $\tilde{\varphi} > \varphi$. If agent i buys the asset and sells contract φ, then any agent $j > i$ buys the asset and sells contract $\tilde{\varphi} > \varphi$

- ***Lemma 3.3*** *Each agent buying (selling) contracts, buys (sells) only one type of contract φ.*

Lemmas 3.1 and 3.3 imply that identities of buyers and sellers are one-to-one functions of contracts. Lemma 3.2 implies these functions are strictly increasing. In the following I provide proofs for these Lemmas.

Lemma 3.1: Proof..
Suppose contract $\varphi \in (\underline{c}, \overline{c})$ is bought by i_1 and i_2, $i_1 \neq i_2$. Then optimality of contracts implies:

$$\frac{1 - F_{i_1}(\varphi)}{R_c(i_1, \varphi)} = q'(\varphi) = \frac{1 - F_{i_2}(\varphi)}{R_c(i_2, \varphi)}$$

Which due to Lemma B contradicts $i_1 \neq i_2$ and $\varphi \in (\underline{c}, \overline{c})$.

Suppose contract $\varphi \in (\underline{c}, \overline{c})$ is sold by j_1 and j_2, $j_1 \neq j_2$. Then optimality of contracts implies:

$$\frac{1 - F_{j_1}(\varphi)}{R_y(i_1, \varphi)} = q'(\varphi) = \frac{1 - F_{j_2}(\varphi)}{R_y(i_2, \varphi)}$$

Which due to Lemma B contradicts $j_1 \neq j_2$ and $\varphi \in (\underline{c}, \overline{c})$.

I am not considering $\varphi \geq \overline{c}$ and $\varphi \leq \underline{c}$ as former contracts provide zero return for sellers and can not be traded in equilibrium, and latter are not traded in equilibrium by Theorem 2.

Lemma 3.2: Proof..
For the first part, assume $j > i$, i buys contract φ and j buys contract $\tilde{\varphi}$. Lemma 3.1 rules out the case $\tilde{\varphi} = \varphi$. Suppose $\tilde{\varphi} < \varphi$. Let's show that in this case j would prefer to buy contract φ: $R_c(j, \varphi) > R_c(j, \tilde{\varphi})$.

$$R_c(j, \varphi) > R_c(j, \tilde{\varphi}) \iff$$

$$\frac{\int_{\underline{c}}^{\varphi}(1 - F_j(y))dy + \underline{c}}{q(\varphi)} > \frac{\int_{\underline{c}}^{\tilde{\varphi}}(1 - F_j(y))dy + \underline{c}}{q(\tilde{\varphi})} \iff$$

$$\underbrace{\int_{\underline{c}}^{\varphi}(1 - F_j(y))dy}_{A_j} > \underbrace{\frac{q(\varphi)}{q(\tilde{\varphi})}}_{\alpha} \underbrace{\int_{\underline{c}}^{\tilde{\varphi}}(1 - F_j(y))dy}_{B_j} + \underbrace{\underline{c}\left[\frac{q(\varphi)}{q(\tilde{\varphi})} - 1\right]}_{\gamma}$$

In other words, one has to show that $A_j > \alpha B_j + \gamma$*, given that* $A_i \geq \alpha B_i + \gamma$ *(since* $R_c(i,\varphi) \geq R_c(i,\tilde{\varphi})$*) and* $j > i$*. It suffices to show that* $\frac{A_j}{B_j} > \frac{A_i}{B_i}$*, since then* $A_j > \frac{A_i}{B_i} B_j \geq \frac{\alpha B_i + \gamma}{B_i} B_j = \alpha B_j + \gamma \frac{B_j}{B_i} > \alpha B_j + \gamma$*, where the last inequality follows from Lemma A.*

$$\frac{A_j}{B_j} = \frac{\int_{\underline{c}}^{\varphi}(1-F_j(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy} = 1 + \frac{\int_{\tilde{\varphi}}^{\varphi}(1-F_j(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_j(y))dy} = 1 + \frac{\int_{\tilde{\varphi}}^{\varphi} \frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}} \frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}$$

$$> 1 + \frac{\frac{1-F_j(\tilde{\varphi})}{1-F_i(\tilde{\varphi})}\int_{\tilde{\varphi}}^{\varphi}(1-F_i(y))dy}{\frac{1-F_j(\tilde{\varphi})}{1-F_i(\tilde{\varphi})}\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} = \frac{\int_{\underline{c}}^{\varphi}(1-F_i(y))dy}{\int_{\underline{c}}^{\tilde{\varphi}}(1-F_i(y))dy} = \frac{A_i}{B_i}$$

For the second part, assume agent i *buys the asset and sells contract* φ*. Consider* $j > i$*. By Lemma 1.2,* j *also buys the asset and leverages with some contract* $\tilde{\varphi}$*. Both* φ *and* $\tilde{\varphi}$ *must be strictly smaller than* $\bar{c}$*, as otherwise* i*'s and* j*'s returns are zero. By Lemma 3.1,* $\tilde{\varphi} = \varphi$ *is ruled out. Suppose* $\tilde{\varphi} < \varphi$*. Let's show that* $R_y(j,\varphi) > R_y(j,\tilde{\varphi})$*.*

$$R_y(j,\varphi) > R_y(j,\tilde{\varphi}) \iff \frac{\int_{\tilde{\varphi}}^{\bar{c}}(1-F_j(y))dy}{\int_{\varphi}^{\bar{c}}(1-F_j(y))dy} < \frac{p-q(\tilde{\varphi})}{p-q(\varphi)}$$

Now consider the left-hand side:

$$\frac{\int_{\tilde{\varphi}}^{\bar{c}}(1-F_j(y))dy}{\int_{\varphi}^{\bar{c}}(1-F_j(y))dy} = 1 + \frac{\int_{\tilde{\varphi}}^{\varphi} \frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}{\int_{\varphi}^{\bar{c}} \frac{1-F_j(y)}{1-F_i(y)}(1-F_i(y))dy}$$

$$< 1 + \frac{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\tilde{\varphi}}^{\varphi}(1-F_i(y))dy}{\frac{1-F_j(\varphi)}{1-F_i(\varphi)}\int_{\varphi}^{\bar{c}}(1-F_i(y))dy} = \frac{\int_{\tilde{\varphi}}^{\bar{c}}(1-F_i(y))dy}{\int_{\varphi}^{\bar{c}}(1-F_i(y))dy} \leq \frac{p-q(\tilde{\varphi})}{p-q(\varphi)}$$

where the first inequality is due to assumption A1 and the last one follows from the fact that i prefers to sell contract φ*.*

Lemma 3.3: Proof..
Suppose i *buys risky contracts* φ_0 *and* φ_1 *such that* $\varphi_0 < \varphi_1$*. It must be the case that* $R_c(i,\varphi_0) = R_c(i,\varphi_1)$*. These contracts are sold by agents* j_0 *and* j_1*, and by Lemma 3.2,* $j_1 \geq j_0 > i$*. Suppose* $j_1 = j_0$ *and thus* j_0 *is indifferent between the two contracts:* $R_y(j_0,\varphi_0) = R_y(j_0,\varphi_1)$*. Optimality conditions then imply:*

$$\frac{1-F_i(\varphi_0)}{R_c(i,\varphi_0)} = q'(\varphi_0) = \frac{1-F_{j_0}(\varphi_0)}{R_y(j_0,\varphi_0)}$$

$$\frac{1-F_i(\varphi_1)}{R_c(i,\varphi_0)} = \frac{1-F_i(\varphi_1)}{R_c(i,\varphi_1)} = q'(\varphi_1) = \frac{1-F_{j_0}(\varphi_1)}{R_y(j_0,\varphi_1)} = \frac{1-F_{j_0}(\varphi_1)}{R_y(j_0,\varphi_0)}$$

and it follows that:

$$\frac{1 - F_{j_0}(\varphi_0)}{1 - F_i(\varphi_0)} = \frac{1 - F_{j_0}(\varphi_1)}{1 - F_i(\varphi_1)}$$

which contradicts the hazard-rate order property. Thus $j_1 > j_0$. *Consider agent* j *such that* $j_0 < j < j_1$, *who sells contract* φ *such that, by Lemma 3.2,* $\varphi_0 < \varphi < \varphi_1$. *Again, by Lemma 3.2, it must be the case that this contract is bought by agent* i. *Since this has to hold for any* $j \in (j_0, j_1)$, *there is a positive mass of contract sellers* $[j_0, j_1]$ *trading with only one (measure zero) buyer* i, *which violates market clearing.*

The proof of the statement that each seller of risky contracts sells one contract only is analogous.

Lemma 2: Proof..
Theorem 3 establishes that there exist one-to-one mappings from the set of traded contracts onto the sets of contracts buyers and sellers. By Theorem 1, the sets of contract buyers and sellers are given by intervals $[i^*, j^*]$ *and* $[j^*, 1]$. *It remains to show that these mappings are continuous, and then it would follow that the set of traded contracts is also an interval.*

Denote by $\varphi(i)$ *the one-to-one mapping from the set of contract buyers into the set of traded contracts. The function* $\varphi(i)$ *is defined implicitly by optimality conditions:*

$$G(i, \varphi) := \frac{1 - F_i(\varphi)}{R_c(i, \varphi)} - q'(\varphi) = 0$$

$G(i, \varphi)$ *is continuous due to the assumptions on differentiability of* $F_i(\varphi)$ *with respect to both* i *and* φ, *and differentiability of* $q(\varphi)$. *For the same reason,* $G'_i(i, \varphi)$ *is continuous. This means that implicit function* $\varphi(i)$ *is continuous for all* $i \in [i^*, j^*]$ *and thus the set of traded contracts is an interval* $[\varphi(i^*), \varphi(j^*)]$, *since* $\varphi(i)$ *is a strictly increasing function.*

Curriculum vitæ

2015–2021	University of Mannheim *Ph.D. in Economics*
2016–2017, 2020	Yale University *Visiting Ph.D. Student*
2013–2015	University of Bonn *M.Sc. in Economics*
2009–2013	Lomonosov Moscow State University *B.Sc. in Economics*
2012	University of Münster *Exchange Student*