Low Inflation: Potential Causes, Effects and Solutions

Christopher David Cotton

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019
ABSTRACT
Low Inflation: Potential Causes, Effects and Solutions
Christopher David Cotton

My dissertation focuses upon low inflation. Many developed countries, especially Japan and the Eurozone, have recently experienced prolonged periods of below-target inflation. This has been blamed for many economic ills including worsening the Great Recession and generating a slow recovery, making monetary policy ineffective and leading to lower labor market flexibility. I study what has caused low inflation, its potential effects and how it could be prevented.

In Chapter 1, I look at how effective raising the inflation target would be in mitigating the problems of low inflation. Many economists have proposed raising the inflation target to reduce the probability of hitting the zero lower bound (ZLB). It is both widely assumed and a feature of standard models that raising the inflation target does not impact the equilibrium real rate. I demonstrate that once heterogeneity is introduced, raising the inflation target causes the equilibrium real rate to fall in the New Keynesian model. This implies that raising the inflation target will increase the nominal interest rate by less than expected and thus will be less effective in reducing the probability of hitting the ZLB. The channel is that a rise in the inflation target lowers the average markup by price rigidities and a fall in the average markup lowers the equilibrium real rate by household heterogeneity which could come from overlapping generations or idiosyncratic labor shocks. Raising the inflation target from 2% to 4% lowers the equilibrium real rate by 0.38 percentage points in my baseline calibration. I also analyse the optimal inflation level and provide empirical evidence in support of the model mechanism.

In Chapter 2, I study to what degree the recent fall in inflation can explain the rise in
firm profitability which has been blamed for a rise in inequality. A theoretical relationship between inflation and profitability is known to exist. I investigate the degree to which the recent fall in inflation can explain the rise in firm profitability. My three primary findings are: 1. The negative relationship between inflation and profitability does not hinge upon the Calvo assumption. Raising inflation significantly lowers profitability under all common price rigidities. The relationship can actually be significantly stronger under menu costs. 2. A rise in the degree to which firms discount the future magnifies the effect; a rise in elasticity of substitution can increase or decrease the effect depending upon the price rigidity. 3. The profit share has risen by around 3.5p.p. since the 1990s. In a richer model with firm heterogeneity, the recent fall in inflation is estimated to explain 14% of the rise. This can increase to 29% if firms are allowed to discount the future by more in line with estimates from the finance literature. I also provide empirical evidence for the negative relationship between inflation and firm profits.

In Chapter 3, I examine whether behavioral features can help to explain why some countries have persistently experienced low inflation at the zero lower bound. Economists are keen to introduce behavioral assumptions into modern macroeconomic models. A popular framework for doing so is sparse dynamic programming, which assumes that agents partly base their expectations upon a default model which is typically the steady state. This means agents’ expectations will be wrong if there are long-run deviations from the default model and assumes agents can compute the default. I introduce an alternative form of sparse dynamic programming which tackles these problems by allowing for long-run updating to the behavioral part of agents’ expectations. I apply this to derive a long-run behavioral New Keynesian model. Within this model, fixed interest rates yield indeterminacy and the costs of remaining at the zero lower bound are unbounded. These results are very different to a behavioral New Keynesian model based upon standard sparse dynamic programming, which can yield determinacy under fixed interest rates and bounded costs of the zero lower bound.
# Contents

List of Figures vi

List of Tables viii

## 1 The Inflation Target and the Equilibrium Real Rate 1

1.1 Introduction ................................................. 2

1.2 Intuition through a Simplified Model ................................................. 8

1.2.1 Relationship between the Inflation Level and the Markup ................................................. 8

1.2.2 Relationship between the Markup and the Equilibrium Real Rate ................................................. 10

1.3 Model .......................................................... 19

1.3.1 Households ................................................. 20

1.3.2 Firms ..................................................... 25

1.3.3 Monetary and Fiscal Policy ................................................. 30

1.3.4 Other Conditions ................................................. 30

1.4 Model Solution and Calibration ................................................. 31

1.4.1 Full Conditions ................................................. 31

1.4.2 Steady State ................................................. 32

1.4.3 Shocks and Log-linearized Conditions ................................................. 32

1.4.4 Policy Functions and Simulation ................................................. 35

1.4.5 Calibration ................................................. 37

1.5 Impact of Raising the Inflation Target ................................................. 38
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.2 Model Details</td>
<td>116</td>
</tr>
<tr>
<td>A.2.1 Firms</td>
<td>116</td>
</tr>
<tr>
<td>A.3 Steady State Details</td>
<td>121</td>
</tr>
<tr>
<td>A.3.1 Markup and Inflation Target</td>
<td>121</td>
</tr>
<tr>
<td>A.3.2 Relative Asset Supply</td>
<td>123</td>
</tr>
<tr>
<td>A.3.3 Relative Asset Demand</td>
<td>124</td>
</tr>
<tr>
<td>A.4 Generalized OLG Theory</td>
<td>128</td>
</tr>
<tr>
<td>A.5 Calibration Details</td>
<td>130</td>
</tr>
<tr>
<td>A.5.1 Endogenous Labor Supply</td>
<td>130</td>
</tr>
<tr>
<td>A.6 Results Details</td>
<td>131</td>
</tr>
<tr>
<td>A.7 Idiosyncratic Labor Model</td>
<td>132</td>
</tr>
<tr>
<td>A.7.1 Households</td>
<td>132</td>
</tr>
<tr>
<td>A.7.2 Other Computations</td>
<td>135</td>
</tr>
<tr>
<td>A.8 Empirics Robustness</td>
<td>135</td>
</tr>
<tr>
<td>B Appendix to Chapter 2</td>
<td>139</td>
</tr>
<tr>
<td>B.1 Introduction Details</td>
<td>140</td>
</tr>
<tr>
<td>B.2 Simple Model Additional Details</td>
<td>141</td>
</tr>
<tr>
<td>B.2.1 Log Linearization of Calvo</td>
<td>141</td>
</tr>
<tr>
<td>B.2.2 Log Linearization of Rotemberg</td>
<td>142</td>
</tr>
<tr>
<td>B.3 Determinants of Long Run Relationship Additional Details</td>
<td>143</td>
</tr>
<tr>
<td>B.3.1 Proof of theorem 1</td>
<td>143</td>
</tr>
<tr>
<td>B.3.2 Menu Costs Additional Figures</td>
<td>144</td>
</tr>
<tr>
<td>B.3.3 Degree of Monetary Non-Neutrality</td>
<td>144</td>
</tr>
<tr>
<td>B.4 Extended Model Details</td>
<td>146</td>
</tr>
<tr>
<td>B.4.1 Multi-Sector Calvo Steady State Derivation</td>
<td>146</td>
</tr>
<tr>
<td>B.5 Empirics Robustness</td>
<td>150</td>
</tr>
</tbody>
</table>
C Appendix to Chapter 3

C.1 Model Setup Details .................................................... 154
  C.1.1 Firm Details ....................................................... 154

C.2 Determinacy ................................................................. 158
  C.2.1 Proof of Theorem 5 ............................................... 158
  C.2.2 Proof of Theorem 6 ............................................... 159

C.3 Costs of Zero Lower Bound ........................................... 162
  C.3.1 Costs of Zero Lower Bound in Standard Behavioral New Keynesian Model ............................................. 162
  C.3.2 Costs of Zero Lower Bound in Long-Term Behavioral New Keynesian Model ............................................. 163
List of Figures

1.1 Relative Asset Supply under a Fall in the Markup ................................. 13
1.2 Equilibrium under a Fall in the Markup: 1. Representative Agent ............... 15
1.3 Equilibrium under a Fall in the Markup: 2. OLG .................................. 17
1.4 Equilibrium under a Fall in the Markup: 3. Idiosyncratic Labor .................. 19
1.5 Baseline Calibration: Asset Supply and Demand .................................. 39
1.6 Baseline Calibration: Consumption Path ........................................... 40
1.7 IES = 0.1: Asset Supply and Demand .............................................. 42
1.8 OLG with Idiosyncratic Labor Shocks: Asset Supply and Demand ............. 44
1.9 Welfare under Baseline Calibration .................................................... 46
1.10 Probability of Hitting the ZLB under Baseline Calibration ...................... 46
1.11 Welfare under Larger Shocks .......................................................... 47
1.12 Probability of Hitting the ZLB under Larger Shocks ............................. 47

2.1 Relationship Between Inflation and the Profit Share under Alternative Forms of Price Rigidity ................................................................. 65
2.2 Change in Profit Share (p.p.) when Inflation Falls from 2% to 0% Under Different Degrees of Firm Discounting ............................................. 68
2.3 Change in Profit Share (p.p.) when Inflation Falls from 2% to 0% Under Different Elasticities of Substitution Between Goods .............................. 69
2.4 Long Run Relationship between Inflation and the Profit Share allowing for Firm Heterogeneity ................................................................. 73
List of Tables

1.1 Policy Experiment with Default Calibration ............................... 41
1.2 Impact of Changing Intertemporal Elasticity of Substitution on Results of Policy Experiment ........................................... 42
1.3 Impact of Changing Elasticity of Labor Supply on Results of Policy Experiment ...................................................... 43
1.4 Empirical Estimates of Relationship between Long-Run Inflation and the Equilibrium Real Rate ............................... 50

2.1 Comparison of Recent Eras of Inflation ........................................ 56
2.2 Relationship between Inflation and the Profit Share under Alternative Price Rigidities Summary Table .......................... 64
2.3 Relationship between Inflation and the Profit Share under Firm Heterogeneity Summary Table ............................. 73
2.4 Empirical Estimates of Relationship between Long Run Inflation and Long Run Markup ............................................. 76

A.1 Impact of Changing Firm Discounting on the Policy Experiment Results ......................................................... 132
A.2 Relationship between Inflation and the Real Rate OECD Original Members ...................................................... 136
A.3 Relationship between Inflation and the Real Rate Low Inflation ................................................................. 136
A.4 Relationship between Inflation and the Real Rate Pre-2000 ................................................................. 137
A.5 Relationship between Inflation and the Real Rate Post-2000 ................................................................. 137
I am indebted to my advisers, Jón Steinsson and Michael Woodford, for their guidance and support in completing my dissertation and throughout my PhD. I am also grateful to other faculty members who have helped me during my time at Columbia University. I especially thank Hassan Afrouzi, Jennifer La’O, Emi Nakamura, Serena Ng and Martin Uribe.

I completed part of this project during a Dissertation Fellowship at the Federal Reserve Board. This experience greatly helped me develop the first chapter of my dissertation. I am particularly grateful to the members of the Monetary Studies section who hosted me and, in particular, Etienne Gagnon and Benjamin Johannsen.

I also wish to thank my friends, especially Ashna Arora, Daniel Rappoport and Scott Weiner, who provided great advice. Finally, I wish to thank Jason and my parents for their support and sense of perspective.
Chapter 1

The Inflation Target and the Equilibrium Real Rate
1.1 Introduction

Many economists have proposed raising the inflation target to reduce the probability of hitting the zero lower bound (ZLB). Nearly all developed countries were constrained by the ZLB during the financial crisis. Moreover, it is widely believed that average real interest rates have fallen. This implies that average nominal interest rates will be lower going forward. Consequently, there has been a re-evaluation of the risk that central banks will hit the ZLB. Hitting the bound is bad for economic outcomes because central banks have less room to lower nominal interest rates and stimulate the economy during bad times. Therefore many economists (including Blanchard et al. (2010), Ball (2014), Krugman (2014)) have proposed raising the inflation target from the standard objective of 2% to 4% claiming this will raise average nominal interest rates and thus reduce the probability of hitting the ZLB.

It is widely assumed that raising the inflation target will not affect the equilibrium real rate. The equilibrium real (nominal) rate is the real (nominal) interest rate on short-term safe assets when there are no shocks. Standard macroeconomic models very commonly assume flexible prices and/or a representative agent. With either of these assumptions, the equilibrium real rate is unaffected by changing average inflation. This is also a historic concept introduced by Fisher (1907) and is often taken for granted within policy discussions. For example, Ball (2014) states that the long run level of the real interest rate is “independent of monetary policy”. Thus, it is widely believed that raising the inflation target by 2p.p. will have no impact upon the equilibrium real rate and will therefore raise the equilibrium nominal rate by a corresponding 2p.p.

My primary contribution is to demonstrate a new channel by which raising the inflation target will lower the equilibrium real rate. Once I account for household heterogeneity
(through either overlapping generations or idiosyncratic risk) within the standard New Keynesian model, I find that raising the inflation target lowers the equilibrium real rate. This implies that a rise in the inflation target will raise the average nominal interest rate by less than expected. Since nominal interest rates will rise by less than expected, raising the inflation target will reduce the probability of avoiding the ZLB by less than is commonly believed. The channel has two stages. Firstly, price rigidities imply that a rise in the inflation target lowers the markup. Secondly, household heterogeneity implies that a fall in the markup lowers the equilibrium real rate.

The first part of the channel is a standard, albeit often overlooked, feature of New Keynesian models. A firm’s markup is just the ratio of its price to its nominal marginal cost. When firms set their prices infrequently, a higher average inflation level has two opposing impacts upon average markups. Firstly, higher inflation means that when a firm does not reset its price then its markup falls by relatively more since with higher inflation nominal marginal costs rise relatively more quickly. Secondly, firms observe that their markups fall more quickly and therefore set their markup to be higher when they do get to reset their prices. It can be shown that with no discounting these two effects cancel out and thus average markups are unchanged by raising average inflation. However, with discounting, the first effect dominates since firms care more about making profits in the current period and so do not want to set their current markup to be very high when they reset their price. Therefore, a rise in average inflation lowers the average markup. I demonstrate in Cotton (2019) that the negative relationship between trend inflation and firm markups has significant effects under all standard price rigidities.

The second part of the channel is that once you allow for household heterogeneity a fall in the markup lowers the equilibrium real rate. Taking the example of heterogeneity through overlapping generations (OLG): A fall in the markup lowers firm profits and thus reduces the value of shares and of overall savings. A fall in the amount of savings ceteris paribus lowers the consumption of the old relative to the young. This means the old have higher
marginal utility from consuming than the young. Thus, there is greater competition among young people to save for when they are old and so the price of savings rises. As the price of savings rises, the return on savings (the equilibrium real rate) falls. To my knowledge, this part of the channel has not been covered in the literature.

This contrasts with a representative agent New Keynesian model where a fall in the markup has no impact on the equilibrium real rate. In this case, a rise in inflation still lowers the markup which in turn lowers firm profits and thus reduces the value of shares. However, within a representative agent framework, the agent’s consumption path does not depend upon average household savings since it does not matter what level of savings the agent chooses to hold during their infinitely long life. Instead, without shocks, they just set their level of consumption to be the same over time and thus the equilibrium real rate is purely determined by the agent’s patience.

I estimate the impact of the channel through a fully calibrated model. I study the effect of raising the inflation target within a model with standard New Keynesian features and a fully calibrated life cycle framework. Within the baseline calibration, I find the equilibrium real rate falls by 0.38 p.p. when the inflation target is raised from 2% to 4%. This implies that average nominal interest rates rise by 1.62 p.p. as opposed to the 2 p.p. that would typically be expected and be found within standard models. Thus, raising the inflation target mitigates the probability of hitting the ZLB by less than expected. When I reduce the intertemporal elasticity of substitution from 0.5 to 0.1, I find the fall in the equilibrium real rate increases to 0.67 p.p.

I compute the optimal change in inflation in response to a fall in the equilibrium real rate as well as the optimal level of inflation more generally. To assess the optimal inflation target, I find the welfare of the simulated path of the economy of my model under different inflation targets taking into account the ZLB with calibrated shocks. Much of the motivation for raising the inflation target is based upon the suggestion that the equilibrium real rate is

---

2Recent research by Best et al. (2018) suggests an intertemporal elasticity of substitution of 0.1.
has fallen. I assess how much the inflation target increases when the equilibrium real rate falls by 2p.p. In my baseline calibration, I find this increases the optimal inflation target by 0.3p.p. When I allow for larger shocks which increase the probability of hitting the ZLB, I find the increase in the optimal inflation target is 0.6p.p. I also analyze the level of the optimal inflation target. The optimal inflation target is always around 1p.p. This is similar to Coibion et al. (2012) who assess the optimal inflation target in a representative agent model. Therefore, the benefits of avoiding the ZLB appear to be dominated by the welfare costs of price dispersion even for relatively low inflation targets.

I also provide empirical evidence for my mechanism. I show there is a negative empirical relationship between long-run inflation and the equilibrium real rate which supports my hypothesized channel. In recent years, inflation and the real interest rate have both fallen across developed countries. This would contradict my channel if the fall in inflation was the only change that could have driven real interest rates lower. However, many factors have been proposed that have lowered real interest rates for other reasons across developed countries. Indeed, it is puzzling that real interest rates have not fallen by more. Gagnon (2009) argue that demographic factors alone can explain the fall in the equilibrium real rate while Eggertsson et al. (2017) argue that real rates should be much lower. I take this into account in my empirical analysis by looking at panel data regressions of the real rate on long-run inflation controlling for country and time fixed effects in OECD countries. The time fixed effects allow me to control for any common change in real rates across countries. I find a 1p.p. rise in long-run inflation lowers the equilibrium real rate by 0.61p.p.

There is a historical literature that looks at the impact of inflation on the equilibrium real rate through non-interest paying money balances but it may be less relevant today. Mundell (1963) and Tobin (1965) argued that when inflation rises, it becomes costlier to hold money.

---

3 My channel would predict a rise in the equilibrium real rate when inflation falls.

4 Stories include: demographic changes (Carvalho et al. [2016], Gagnon et al. [2016]), global savings glut (Caballero and Farhi [2017]), secular stagnation (Eggertsson and Mehrotra [2014]), low productivity growth (Yi and Zhang [2017]), high inequality (Lancastre [2018]).
so agents save more in capital, leading to a fall in the equilibrium real rate. A key assumption of this literature is that money does not pay interest. This has two important implications. Since money does not pay interest, agents need some other incentive to hold money such as the assumption of a cash-in-advance constraint or money-in-utility. Secondly, most central banks in developed countries have now shifted to a framework where they control nominal interest rates by paying interest on reserves in which case a rise in inflation will lead to higher interest on reserves and thus no portfolio shift to capital away from money. Therefore, this literature appears less relevant to modern central banking. My proposed channel is very different because it does not rely upon money holdings in any way.

My model relates to several interesting literatures on heterogeneous agent models: 1. the allocation of profits, 2. redistributional effects of monetary policy, 3. optimal monetary policy with heterogeneous agents. Unlike many heterogeneous agent models, I allow for the endogenous allocation of profits. Most heterogeneous agent models exogenously allocate profits i.e. certain agents are assigned to receive profits. For instance, Werning (2015) considers how these exogenous profit allocations impact the marginal propensity to consume and related implications. I instead consider the case where agents only receive profits by owning shares in firms which get traded each period. Thus, it is an endogenous feature of my model that old people naturally consume less as a result of the fall in the markup.

Raising the inflation target within a heterogeneous agent model can generate interesting long-run distributional effects. Raising the inflation target can have short-term redistributional effects which hurt savers and benefit borrowers by lowering the value of nominal assets. Doepke et al. (2015) consider these short-term redistributional effects in detail. My paper implies that there can actually be long-run redistributional effects as well. A rise in the inflation target reduces profits and thus the value of shares and total savings. This implies that old people, who rely upon savings, consume relatively less and young people consume

---

5There are many other papers in this literature. For instance, Stockman (1981) proposed a reverse-Mundell-Tobin effect in which raising inflation raises the equilibrium real rate due to a cash in advance constraint on investment i.e. you need to take money out of the bank a period before investing.
relatively more indefinitely as a result of a rise in the inflation target.

I contribute to the literature on optimal monetary policy in heterogeneous agent models. I investigate optimal monetary policy within a New Keynesian model with OLG features. Lepetit (2017) shows that within a New Keynesian model with perpetual youth, it can be optimal to set a positive inflation target because heterogeneity can imply that private discounting is higher than social discounting. In this case, central banks raise inflation to lower average markups. My paper is quite different because the primary reason central banks want to raise inflation above zero is to avoid hitting the ZLB which Lepetit does not consider.

There is empirical evidence that supports my channel. Other papers have suggested that raising the long-run inflation rate lowers the long-run real interest rate. King and Watson (1997) consider the impact of raising inflation upon the real interest rate and show that an increase in long-run inflation leads to a decrease in the long-run real interest rate regardless of the restrictions imposed in a structural VAR model for US data. They find that a rise in of 1p.p. in long-run inflation lowers the equilibrium real rate by 0.66p.p. Rapach (2003) extends the analysis to 14 countries with a richer structural model. He demonstrates that a rise in long-run inflation leads a fall of between 0.94p.p. and 0.59p.p. in the equilibrium real rate.

In section 1.2 I outline a simple model that captures the key features found in the rest of the paper. I then outline the full model (section 1.3). I discuss the model solution and calibration in section 1.4. I use the full model to analyse how changing the inflation target will impact the equilibrium real rate in section 1.5. I then consider the optimal inflation target in section 1.6. I discuss my supporting empirical results in section 2.6. Section 1.8 concludes.
1.2 Intuition through a Simplified Model

I break the intuition for the channel into two parts. First, it is demonstrated that a rise in inflation lowers the average markup through firms’ pricing decisions. Next, it is shown that a fall in the markup lowers the equilibrium real rate through multiple forms of household heterogeneity.

1.2.1 Relationship between the Inflation Level and the Markup

A firm’s markup, denoted $m_t$, is its current price, $P_t^*$, divided by its nominal marginal cost, $MC_t$:

$$m_t = \frac{P_t^*}{MC_t}$$

Firms’ profits depend upon their markup. If they set their markup too high, they will not make enough sales. If they set it too low, they will make a lot of sales but with too little profit on each sale. When firms have fully flexible prices, they can set their price so that their markup yields the maximum profits each period. In the common case where firms face constant elasticity of demand, the optimal markup is just $\frac{\sigma}{\sigma - 1}$ where $\sigma$ is the CES parameter.

Setting markups is more complex in the case with infrequent price adjustment. When firms can only change their price infrequently, they are no longer able to set the optimal flexible price markup each period. In the case of positive inflation: There will be two important effects. Firstly, if firms do not get to change their price in a period then their markup will fall. This is because their nominal marginal costs ($MC_t$) rise (due to the rise in the price level) while their price ($P_t^*$) remains constant. Secondly, in anticipation that they may not get to change their price in the future and thus their markup will fall, firms will set their markups to be higher than the optimal flexible price markup when they do get to change their price.
The impact of raising inflation on the markup depends upon the degree of discounting. In the case with no discounting, firms will weight their profits equally in current and future periods. This leads to a special case where the markup is unaffected by changing the level of inflation since the two effects on the markup cancel out. However, when firms discount the future, they will weight their current period markup more in their decision-making. This implies that they set a lower markup when they get to change their price and thus that the average markup is lower with positive inflation. As the level of inflation rises, the strength of this effect will increase.

For example, in the case of Calvo pricing: Denote $\sigma$ to be the elasticity of substitution between goods, $\beta$ to be the discount factor, $\lambda$ to be the probability with which firms can update their price each period and $\bar{\Pi}$ to be the steady state level of gross inflation (i.e. $\frac{P_{t+1}}{P_t}$). Then the firms steady state average markup $\bar{m}$ is given by:

$$\bar{m} = \frac{\sigma}{\sigma - 1} \left[ \frac{\bar{\Pi}^{1-\sigma} - (1 - \lambda) \beta}{\bar{\Pi}^{1-\sigma} - (1 - \lambda)} \right] \left[ \frac{\bar{\Pi}^{-\sigma} - (1 - \lambda) \beta}{\bar{\Pi}^{-\sigma} - (1 - \lambda)} \right]$$

(1.1)

When $\beta = 1$ so there is no discounting, equation (1.1) simplifies to give $\bar{m} = \frac{\sigma}{\sigma - 1}$ so raising inflation has no impact upon the markup and firms will set their markup to be at the same level as without price stickiness. However, when $\beta < 1$ so there is discounting, raising inflation always lowers the markup. This is easy to see in the extreme case of full discounting when $\beta = 0$ in which case equation (1.1) simplifies to:

$$\bar{m} = \frac{\sigma}{\sigma - 1} \frac{\bar{\Pi}^{1-\sigma} - (1 - \lambda) \bar{\Pi}}{\bar{\Pi}^{1-\sigma} - (1 - \lambda)}$$

The frequency of price changes does not affect the relationship at low levels of inflation. If the frequency with which firms adjust prices increases, this would reduce the feedback

---

6 The derivations are shown in appendix A.1.1.

7 I prove that when steady state inflation rises the markup falls for all $\beta < 1$ in appendix A.1.1.
from inflation to the markup. However, Gagnon (2009) demonstrates that the frequency with which firms change their price does not appear to vary below annual rates of inflation of 10%. This makes sense because firms are likely to change their price for other reasons (like idiosyncratic demand or costs) than just inflation so the frequency of price changes does not need to change with low inflation.

The negative inflation-markup relationship also holds with price rigidities based upon adjustment costs. The relationship would hold in the case of menu costs (fixed costs of updating prices) or Rotemberg costs (convex adjustment costs of updating prices). The intuition is that firms prefer to pay the cost of updating their price in the future (with positive discounting) so they set a lower markup when inflation rises.

This is a general result. To get a negative relationship between inflation and the markup, I require that firms set their prices infrequently and discount the future. Nakamura and Steinsson (2008) demonstrates that firms have low frequencies of price changes. Jagannathan et al. (2016) demonstrates that firms discount the future significantly. It is also worth stressing that this relationship is present in the representative agent New Keynesian model - nothing here depends upon household heterogeneity.

1.2.2 Relationship between the Markup and the Equilibrium Real Rate

Simple Model I have shown that markup is determined by the level of inflation so I take the markup as given and concentrate upon the real side of a simple model.

Firm’s produce using a linear production function. Therefore, output \( Y_t \) equals labor

---

8 This occurs because firms would set their markup for shorter periods of time on average which means that the markup would fall by less before being changed for a given level of inflation.

9 Here we have effectively assumed that firms do not face price dispersion. This would be true, for example, under Rotemberg Pricing. The case with price dispersion generates exactly the same equations but is a little bit more complicated to derive. The results are shown in appendix A.1.2.
The real marginal cost of firms $MC_t$ will just be the real wage $W_t$:

$$MC_t = W_t \tag{1.3}$$

The markup $m_t$ is just the price divided by the nominal marginal cost which, by definition, equals the inverse of the real marginal cost ($\frac{1}{MC_t}$ so we can rewrite the marginal cost wage (equation (1.3)) relationship as:

$$\frac{1}{m_t} = W_t \tag{1.4}$$

The total real profits $\Omega_t$ of firms will just be their real sales which is just their output minus their costs of labor:

$$\Omega_t = Y_t - W_t L_t$$

$Y_t$ can be substituted out with $L_t$ (equation (1.2)) and then multiply and divide the first term on the RHS by $W_t$ to give:

$$\Omega_t = \left( \frac{1}{W_t} - 1 \right) W_t L_t$$

I then apply the markup-wage relationship (equation (1.4)):

$$\Omega_t = (m_t - 1) \frac{W_t}{P_t} L_t \tag{1.5}$$

**Asset Supply** I break down the solution into the supply and demand for assets. Asset supply is the amount of assets that are available for households to hold. Asset demand is the amount of assets that agents want to hold.

The only asset that agents can save in is shares in firms. The total real value of firms is
denoted by $Z_t$. Therefore, asset supply, denoted $A^s$, is given by:

$$A^s = Z$$  \hspace{1cm} (1.6)

By standard asset pricing, we know that the price of buying $b$ shares $bZ_t$ must equal the next period return on those shares discounted at $r_{t+1}$. The next period return of those shares is the dividends received from profits ($b\Omega_{t+1}$) plus the price the shares are sold for at $t+1$ ($bZ_{t+1}$). Therefore:

$$Z_t = \frac{\Omega_{t+1} + Z_{t+1}}{1 + r_{t+1}}$$  \hspace{1cm} (1.7)

In steady state, we can rewrite equation (1.7) as:

$$\bar{Z} = \frac{\Omega}{\bar{r}}$$  \hspace{1cm} (1.8)

We can substitute the value of shares with profits (equation (1.8)) in the asset supply equation (equation (1.6)):

$$\bar{A}^s = \frac{\Omega}{\bar{r}}$$  \hspace{1cm} (1.9)

We can then substitute out profits using equation (1.5):

$$\bar{A}^s = \bar{m} - \frac{1}{\bar{r}} \bar{W} \bar{L}$$  \hspace{1cm} (1.10)

To make the problem simpler, we define relative assets $a$ which are assets in terms of labor income:

$$a = \frac{A}{WL}$$  \hspace{1cm} (1.11)

The supply of assets (equation (1.10)) can be rewritten in relative terms to get:

$$\bar{a}^s = \frac{\bar{m} - 1}{\bar{r}}$$  \hspace{1cm} (1.12)
Two features can be observed. Firstly, in equation (1.12) a rise in $\bar{r}$ lowers $\bar{a}^s$. This makes sense because higher discounting implies the discounted sum of profits is lower so the value of firms falls. Equation (1.12) is plotted in figure 1.1. The blue curve represents $\bar{a}^s$ with $\bar{m} = 1.3$ and the orange curve represents $\bar{a}^s$ with $\bar{m} = 1.2$. Since raising $\bar{r}$ lowers $\bar{a}^s$, the curve has a downward slope. It may appear strange that the supply curve is downward sloping but this is because the vertical axis is the return on assets. The return on assets is like the inverse of the price of assets (since as the price of assets rises, the return agents make on those assets falls). If the curve was drawn with the price of assets on the vertical axis, it would have the usual upward sloping supply curve.

Secondly, observe that in equation (1.12) a fall in the markup $\bar{m}$ lowers the relative asset supply $\bar{a}^s$ for any real interest rate $\bar{r}$. This makes sense because when the markup falls, the value of firms falls and thus the value of owning shares in firms falls. This can also be seen in figure 1.1. Observe that the fall in the markup shifts the relative asset supply curve left from the blue curve with markup 1.3 to the orange curve with markup 1.2.
Asset Demand: 1. Representative Agent  Next, I consider the shape of the asset demand under three different household structures: 1. Representative agent. 2. Heterogeneity through overlapping generations. 3. Heterogeneity through idiosyncratic labor.

In all standard representative agent problems, we derive an Euler condition of similar form to the following (I assume log utility to keep things very simple):

\[ C_{t+1} = \beta (1 + r_{t+1}) C_t \quad (1.13) \]

A steady state equilibrium requires that a representative agent consumes the same amount over time. If \( C_{t+1} \) is more (less) than \( C_t \) the Euler condition requires that \( 1 + r_{t+1} \) is more (less) than \( \beta \). Therefore, the only way we can have a steady state is when \( \beta (1 + r_{t+1}) \) is stable which requires:

\[ \bar{r} = \frac{1}{\beta} - 1 \quad (1.14) \]

Equation (1.14) is plotted in figure 1.2. Like in figure 1.1, the impact of a fall in the markup is considered. Observe that the asset demand is just a horizontal line since \( \bar{r} \) is always pinned down. Thus, a shift left in the supply of assets lowers the amount of assets held by the household but has no impact upon \( \bar{r} \).

The reason the equilibrium real rate is unchanged is because in steady state the path of consumption of the agent must always be flat i.e. \( C_t = C_{t+1} \) by the Euler condition. Relative asset demand always adjusts to ensure this holds. Therefore, changing the assets held by the household cannot disturb the path of consumption of the agent. Thus, the marginal utility of the agent must always be flat i.e. \( u'(C_t) = u'(C_{t+1}) \) regardless of changes in the supply of assets. The only way the marginal utility of the agent can be flat is if \( \bar{r} \) remains the same over time by the Euler condition.

Asset Demand: 2. Overlapping Generations  Now, household heterogeneity is introduced. The implication in both cases of household heterogeneity that are considered is
that the level of assets does impact the path of the household’s marginal utility over time, meaning that the equilibrium real rate will be impacted by changing the markup.

I first consider a simple overlapping generations model based upon (Diamond, 1965). Every period a new generation is born. Each generation lives for two periods and then dies. The utility of a young agent is given by:

\[
\log(C_{1,t}) + \beta \log(C_{2,t+1})
\]  

(1.15)

Log utility is used for simplicity. Young agents work $L$ unit and devote their income to either consumption $C_{1,t}$ or asset purchases $A_{t+1}$:

\[
C_{1,t} + A_{t+1} = W_t
\]  

(1.16)

Old agents merely consume $C_{2,t+1}$ from their available assets. Their available assets are
their assets from when they were young on which they have earned a return of $r_{t+1}$:

$$C_{2,t+1} = (1 + r_{t+1})A_{t+1}$$  \hspace{1cm} (1.17)

The amount the young save can be solved for by inputting equations (1.16) and (1.17) into equation (1.15) and then taking first-order conditions. This yields:

$$A_{t+1} = \frac{\beta}{1 + \beta} W_t L$$  \hspace{1cm} (1.18)

So agents save some constant fraction of their income each period. It is simpler to rewrite the agent’s demand for assets in relative terms so equation (1.18) is divided by labor income and also written in steady state terms to yield:

$$\bar{a}^d = \frac{\beta}{1 + \beta}$$  \hspace{1cm} (1.19)

In this case, the demand for savings is perfectly inelastic to changes in $\bar{r}$. This is something of a special case (due to log utility and only having two periods). In the full model, demand for relative assets is not perfectly inelastic. However, the generation structure in the full model is still such that the elasticity of demand is not perfectly elastic and thus the real interest rate changes in response to a shift left in the demand for assets.

Equation (1.19) is plotted in figure 1.3 where the impact of a fall in the markup is considered (as in figure 1.1). Observe that the asset demand is just a vertical line since $\bar{a}^d$ is fixed. Thus, a shift left in the supply of assets lowers $\bar{r}$ but has no impact upon the amount of assets demanded by the agent. The relative asset supply on this graph looks a bit different to previous asset demand/supply graphs since each period represents a generation and lasts for 25 – 30 years so it is necessary to rescale the curves to get back to an annual basis.\textsuperscript{10}

This is effectively the opposite to the representative agent case. The reason the impact is

\textsuperscript{10}The non-annualized case is shown in appendix A.1.3.
so different is that a fall in the amount of savings held by the consumer affects the marginal utility of consumption of the young compared to the old. When assets fall, the old consume less relative to the young ceteris paribus. Thus, old people have a relatively higher marginal utility. Therefore, the price of assets rises since young agents are keener to save assets for when they are old. Consequently, the equilibrium real rate falls.

Asset Demand: 3. Idiosyncratic Labor  Within this paper, household heterogeneity is primarily introduced through overlapping generations. However, an extension with idiosyncratic labor is considered and it is worthwhile to demonstrate that a similar intuition explains why the channel holds in this case.

There are many agents, each denoted with subscript $i$. Agents live forever and maximise their lifetime utility:

$$\max_{\sum_{t=0}^{\infty}} \mathbb{E}_0[\beta^t u(C_{i,t})]$$
Agents receive a wage $W$ from the amount they work $L_{i,t}$, which varies over time and across agents, and some real return $r$ on assets $A_{i,t}$. Agents spend their money on consumption $C_{i,t}$ and assets for the next period. Note that there are no aggregate shocks hence why $W, r$ have no time subscripts. Their budget constraint is:

$$C_{i,t} + A_{i,t+1} = (1 + r)A_{i,t} + WL_{i,t}$$

A key additional feature is that agents face some borrowing constraint, which is set to be 0, and this limits the amount they may borrow each period:

$$A_{i,t+1} \geq 0$$

This problem can be solved by value function iteration. Ultimately, we get effectively the same solution as Aiyagari (1994)\(^{11}\) The asset demand $\bar{a}$ can then be computed for any equilibrium real rate $\bar{r}$.

$\bar{a}$ is plotted in figure 1.4 where a fall in the markup is considered (as in figure 1.1). A shift left in the asset supply due to a fall in the markup leads to a fall in relative assets and a fall in the equilibrium real rate.

The result of lowering the markup is different to the representative agent case because a fall in assets lowers the marginal utility in the next period by more than the current period since it means that more agents will face a binding borrowing constraint in the next period.

\(^{11}\) It is necessary to make a minor change from Aiyagari which is to rewrite the problem using relative assets (this has no substantive impact upon the results however):

$$\max \mathbb{E}_0 \left[ \sum_{t=0}^{\infty} \beta^t u(c_{i,t}) \right]$$

s.t.

$$a_{i,t+1} = (1 + r)a_t + \frac{L_{i,t}}{L_t} - c_{i,t}$$

$$a_{i,t+1} \geq 0$$

18
This means that agents want to save more. In turn, this raises the price of assets and lowers their real return in equilibrium (the equilibrium real rate).

The degree to which a shift left in assets lowers the equilibrium real rate depends upon whether many agents are close the borrowing constraints. When the level of assets is high (low), a fall in assets will increase a little (lot) the number of agents affected by the borrowing constraint so it will raise the demand for savings a little (lot) and thus lower the equilibrium real rate a little (lot). This can be seen in figure 1.4 and is reflected graphically in the steeper relative asset demand when relative assets are low.

1.3 Model

I now introduce the full model which is used to assess the importance of the channel and to conduct welfare analysis.
1.3.1 Households

I start by describing the general overlapping generations framework. Each agent lives for $M$ periods. Agents born in different periods overlap. An agent is denoted by its age in periods so an agent born $i$ periods ago is denoted $i$. Therefore, the $M$ cohorts in any given period are denoted $0, \ldots, M-1$. Each period: new agents are born (cohort 0), the oldest agents from the previous period (cohort $M-1$ at time $t-1$) have died and all other generations mature from cohort $i$ to $i+1$.

The population of the cohort born at time $t$ is defined as $N_t$. The total population is defined as $N_t$ and thus $N_t = \sum_{i=0}^{M-1} N_{t-i}$. It is assumed that the population grows at a constant rate of $n$ so that $N_{t+1} = (1 + n)N_t$. Thus, the total population also grows by $1 + n$ each year.

An agent of cohort $i$ at time $t$ has a budget constraint given by equation (1.20). An agent of cohort $i$ consumes $C_{i,t}$ at time $t$. An agent of cohort $i$ works for $L_{i,t}$. $W_t$ is the real wage paid at time $t$ for each unit of work. An agent can invest in bonds, capital or shares in firms. $B_{i,t}, K_{i,t}$ are respectively the bonds and capital held by agents of cohort $i$ at the start of period $t$ (so they were chosen at $t-1$ when that agent was cohort $i-1$). The bond is in nominal terms and pays interest rate $i_{t-1}$ at time $t$ (denoted with a $t-1$ since the nominal interest rate is chosen at $t-1$). Capital is in real terms and agents get a real return of $r_t$ from selling their capital to the firm at time $t$. $\tilde{\omega}_{i,t}$ is the number of shares of the composite firm that agent $i$ owns at the start of time $t$. The total number of shares issued is 1 so $\tilde{\omega}_{i,t}$ also represents the proportion of the firm owned by an agent of cohort $i$ at time $t$. The price of a share is $\tilde{Z}_t$ and it pays out a proportional amount of the firm’s total profits $\tilde{\Omega}_t$ each period. Assume the agent starts with zero assets so $K_{i,0} = B_{i,0} = \tilde{\omega}_{i,0} = 0$. For ease of notation, real and gross interest rates are also defined $R_t = 1 + r_t$, $I_t = 1 + i_t$.

\[
C_{i,t} + \frac{B_{i+1,t+1}}{P_t} + K_{i+1,t+1} + \tilde{Z}_t\tilde{\omega}_{i+1,t+1} \leq W_t L_{i,t} + I_{t-1} \frac{B_{i,t}}{P_t} + R_t K_{i,t} + (\tilde{\Omega}_t + \tilde{Z}_t)\tilde{\omega}_{i,t} \quad (1.20)
\]
The agent’s lifetime utility function when they are in cohort $k$ is given by equation (1.21). CRRA utility is used (equation (1.22)). Both endogenous and exogenous labor are allowed for. In the exogenous labor case, the labor supply is fixed by each cohort so that $L_{i,t} = L_i \forall t$ and the disutility of labor term $v(L_{i,t})$ does not appear in the utility function. In the endogenous labor case, the disutility of labor is given by equation (1.23) where $\eta$ is the elasticity of labor supply. Bonds are also allowed to have additional utility to the consumer. This is not key to the analysis and is only used to easily adjust the real interest on bonds when the optimal inflation target is considered in section 1.6. $u_b$ (equation (1.24)) is set so that the utility on bonds simplifies to give a fixed wedge between the steady state real interest rate on bonds and the steady state real interest rate on other assets (this is $\xi$ in equation (1.25)).

$$E_t \left[ \sum_{i=k}^{M-1} \beta^{i-k} [u(C_{i,t}) + u_b \left( \frac{B_{i,t}}{P_t} \right) - v_i(L_{i,t})] \right]$$ (1.21)

where:

$$u(C) = \frac{C^{1-\gamma}}{1 - \gamma}$$ (1.22)

$$v_i(L_{i,t}) = \frac{1}{1 + \eta} x_i L_{i,t}^{1+\eta}$$ (1.23)

$$u_b \left( \frac{B_{i,t}}{P_t} \right) = \xi I_{t-1} \frac{B_{i,t}}{P_t} u'(\bar{C}_{i,t})$$ (1.24)

Therefore, an agent of age $k$ faces the following problem:

$$\max_{\{C_{i,t}, B_{i+1,t+1}, \ldots, \bar{C}_{i+1,t+1}\}_{i=k}^{M-1}} E_t \left[ \sum_{i=k}^{M-1} \beta^{i-k} [u(C_{i,t}) + u_b \left( \frac{B_{i,t}}{P_t} \right) - v(L_{i,t})] \right]$$

\footnote{\textsuperscript{12} $u_b$ (equation (1.24)) is set so that the wedge simplifies easily in the Fisher equation. $\xi$ is some constant utility from bonds. If $\xi = 0$, then the standard case without utility on bonds applies. Assuming $\xi > 0$ then the utility from bonds depends upon the nominal return from bonds $I_{t-1} \frac{B_t}{P_t}$ and the marginal utility of consumption. It is assumed that the agent does not take into account how changing their consumption will affect the marginal utility from safe bonds so that the condition simplifies easily. This is why $\bar{C}_{i,t}$ is denoted with a bar.}
s.t. \( \forall i \in k, \ldots, M - 1: \)

\[
C_{i,t+i} + \frac{B_{i+1,t+i+1}}{P_{i+i}} + K_{i+1,t+i+1} + \tilde{Z}_{t+i+1} \tilde{\omega}_{i+1,t+i+1} \\
\leq W_{t+i} L_{i,t+i} + I_{t-1} \frac{B_{i,t+i}}{P_{i+i}} + R_{t+i} K_{i,t+i} + (\hat{\Omega}_{t+i} + \tilde{Z}_{t+i}) \tilde{\omega}_{i,t+i}
\]

\[
B_{M,t+M-k}, K_{M,t+M-k}, \omega_{M,t+M-k} \geq 0
\]

First-order conditions are applied. This yields arbitrage conditions on bonds (equation (1.25)), capital (equation (1.26)) and shares (equation (1.27)). Note that gross inflation is defined in the usual way \( (\Pi_{t+1} = \frac{P_{t+1}}{P_t}) \). Also observe that the only impact of the utility on bonds (equation (1.24)) is to add the constant wedge \( (\xi) \) into equation (1.25).

\[\forall i \in 0, \ldots, M - 2:\]

\[
\omega_{i,t} = \tilde{\omega}_{i,t} N_t \Rightarrow \omega_{i,t} \text{ represents the proportional per capita holdings of an agent of cohort } i \text{ at time } t \text{ of firm shares rather than the aggregate holdings of cohort } i \text{ at } t.
\]

Then define \( \tilde{Z}_t \) to be the price of a per capita share in firms i.e. \( \tilde{Z}_t = \frac{\tilde{Z}_t}{N_t} \) and \( \Omega_t \) to be the profits paid by a per capita share in firms i.e. \( \Omega_t = \frac{\hat{\Omega}_t}{N_t} \). Equations (1.20) and (1.27) become respectively:

\[
Z_t u'(C_{i,t}) = \beta \mathbb{E}_t [u'(C_{i+1,t+1})(1 + n)(\Omega_{t+1} + Z_{t+1})]
\]

With endogenous labor, it is derived \( \forall i \in 0, \ldots, M - 1: \)

\[
W_t u'(C_{i,t}) = v'(L_{i,t})
\]
\[ C_{i,t} + \frac{B_{i+1,t+1}}{P_t} + K_{i+1,t+1} + \frac{Z_t \omega_{i+1,t+1}}{1 + n} \leq W_t L_{i,t} + I_{t-1} \frac{B_{i,t}}{P_t} + (1 + r_t) K_{i,t} + (\Omega_t + Z_t) \omega_{i,t} \quad (1.30) \]

All conditions needed to study the long-run equilibrium have been derived. However, to consider the impact of shocks, it is necessary to make some further adjustments to the household conditions.

Define the amount that agents of cohort \( i \) have available at the start of \( t \) from savings they made in \( t - 1 \) as \( T_{i,t} \) (equation (1.32)). Define the amount that agents of cohort \( i \) save at \( t \) for \( t + 1 \) as \( S^p_{i+1,t} \) (equation (1.31))\(^{13}\)\(^{14}\)

\[ S^p_{i+1,t} = \frac{B_{i+1,t+1}}{P_t} + K_{i+1,t+1} + \frac{Z_t \omega_{i+1,t+1}}{1 + n} \quad (1.31) \]

\[ T_{i,t} = I_{t-1} \frac{B_{i,t}}{P_t} + R_t K_{i,t} + (\Omega_t + Z_t) \omega_{i,t} \quad (1.32) \]

Observe that the budget constraint (equation (1.30)) can be rewritten as:

\[ C_{i,t} + S^p_{i+1,t} \leq W_t L_{i,t} + T_{i,t} \quad (1.33) \]

Define \( T_t \) to be the per capita aggregate savings held at the start of a period \( t \) from savings made at \( t - 1 \) (equation (1.34)). This can be computed this by summing the population-weighted savings held by each cohort \( \sum_{i=0}^{M-1} N_{i-t} T_{i,t} \) divided by the total population \( N_t \).

It is then possible to simplify this slight by rewriting the population structure \( N_{t-i}, N_t \) in terms of \( n \). These steps are shown in equation (1.34).

\[ T_t = \sum_{i=0}^{M-1} \frac{N_{t-i} T_{i,t}}{N_t} = \sum_{i=0}^{M-1} \frac{N_{t-i} T_{i,t}}{N_t} \frac{1}{(1 + n)^i} T_{i,t} = \sum_{i=0}^{M-1} \frac{1}{(1 + n)^i} T_{i,t} \quad (1.34) \]

Using the simplified per capita definition, additional variables are defined: \( S^p_t \) is the per

\(^{13}\)A superscript \( p \) is used to represent the fact that these are savings held by agents at the end of \( t \) (which is different to how capital and \( T_t \) are defined).

\(^{14}\)Note that \( T_{0,t}, S^p_{M,t} = 0 \) which makes sense since agents don’t hold assets when they are born or when they are about to die.
capita savings made at $t$ for $t + 1$ (equation (1.35)); $B_t$ is the per capita bonds held at the start of period $t$; $K_t$ is the per capita capital held at the start of period $t$ (equation (1.37)); $\omega_t$ is the per capita holdings of shares at the start of period $t$ (equation (1.38)).

\[
S^p_t = \frac{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i} S^p_{i+1,t}}{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i}}
\]  
(1.35)

\[
B_t = \frac{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i} B_{i,t}}{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i}}
\]  
(1.36)

\[
K_t = \frac{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i} K_{i,t}}{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i}}
\]  
(1.37)

\[
\omega_t = \frac{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \omega_{i,t}}{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i}}
\]  
(1.38)

Next, recall that the total holdings of shares ($\tilde{\omega}_{i,t}$) in a firm must sum to 1 i.e. $\sum_{i=0}^{M-1} N_{t-i} \tilde{\omega}_{i,t} = 1$. Applying the definition of $\omega_{i,t} = \tilde{\omega}_{i,t} N_t$ implies that the aggregate per capita holdings of shares in a firm $\omega_t$ must also always equal 1. This is shown formulaically in equation (1.39).

\[
\omega_t = \frac{\sum_{i=0}^{M-1} N_{t-i} \omega_{i,t}}{N_t} = \sum_{i=0}^{M-1} N_{t-i} \tilde{\omega}_{i,t} N_t = \sum_{i=0}^{M-1} N_{t-i} \tilde{\omega}_{i,t} = 1
\]  
(1.39)

Inputting equation (1.32) into equation (1.34) and then applying equations (1.36) to (1.39) yields equation (1.40). Equation (1.40) just states that the total assets held at $t$ equal the total return on bonds, capital and shares. Similarly, inputting equation (1.31) into equation (1.35) and then applying equations (1.36) to (1.39) yields equation (1.41). \(^{15}\)

\(^{15}\)To derive equation (1.41), the following steps can be made for bonds, capital and shares:

$$
\frac{\sum_{i=0}^{M-1} N_{t-i} K_{i+1,t+1}}{N_t} = \frac{\sum_{i=1}^{M-1} N_{t-i} K_{i,t+1}}{N_t} = \frac{\sum_{i=0}^{M-1} N_{t-i} K_{i,t+1}}{N_t} = (1 + n) \frac{\sum_{i=0}^{M-1} N_{t-i} K_{i,t+1}}{N_{t+1}} = (1 + n) K_{t+1}
$$

The first equality is just an adjustment of the summation index. The second equality uses $K_{0,t+1} = K_{M,t+1} = 0$ to adjust the summation begin and start points. The third equality adjusts $N_t$. The fourth equality is just a definition.
tion (1.41) just states that the total savings made at \( t \) equals next period capital and bonds plus the value of shares purchased.

\[
T_t = I_{t-1} \frac{B_t}{P_t} + R_t K_t + \Omega_t + Z_t 
\]  
(1.40)

\[
S_t^p = (1 + n) \frac{B_{t+1}}{P_t} + (1 + n)K_{t+1} + Z_t 
\]  
(1.41)

Define the share of savings of each cohort to be \( s_{i,t-1}^p = \frac{S_{i,t-1}^p}{S_t^p} \). The definition of equation (1.35) can then be applied to show that the per capita value of \( s_{i,t}^p \) equals 1 in equation (1.42).

\[
1 = \frac{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i} s_{i,t-1}^p}{\sum_{i=0}^{M-1} \frac{1}{(1+n)^i}} 
\]  
(1.42)

Next, set \( T_{i,t} = s_{i,t-1}^p T_t \). This implies that the amount of total assets that a cohort holds at time \( t \) is proportional to the share of saving they did at time \( t - 1 \). Equation (1.33) then can be rewritten as:

\[
C_{i,t} + s_{i+1,t}^p S_{t+1} = W_t L_{i,t} + s_{i,t-1}^p T_t 
\]  
(1.43)

### 1.3.2 Firms

**Final Goods Firm** There is a single competitive final goods firm which aggregates goods in different industries to produce a final good. There are \( J \) industries in total, denoted \( 1, \ldots, J \). The final goods firm has CES production and each industry has a weight \( a_j \) in

\(^{16}\) We will consider a first order perturbation. A first order perturbation means that agents only care about the expected return and do not care about about risk. Therefore, all cohorts are indifferent between holding equivalently valued capital, bonds or shares since they all give a real expected return of \( E_t[R_{t+1}] \). This would also hold without a first order perturbation in a purely deterministic model without any risk. Thus, although the savings of each cohort is known, how savings are comprised is not known i.e. \( \frac{B_{i+1,t+1}}{K_{i+1,t+1}} + K_{i+1,t+1} + Z_{i+1,t} \omega_{i+1,t+1} \) is known but not \( \frac{B_{i+1,t+1}}{K_{i+1,t+1}} \) or \( K_{i+1,t+1} \) or \( \omega_{i+1,t+1} \). This does not matter when there is no risk since these assets will always return the same by arbitrage. However, it does matter when there is risk since if profits fall, agents who hold relatively more shares suffer. The model is kept simple by effectively assuming that agents hold proportional amounts of bonds, capital and shares. This avoids complications where shocks lead to unexpected redistribution.
production:

\[
\left( \sum_{j=1}^{J} a_j \frac{1}{\sigma_j} Y_{j,t}^{\frac{\sigma_j-1}{\sigma_j}} dj \right)^{\frac{\sigma_j}{\sigma_j-1}} = Y_t
\]

Therefore, the final goods firm has the usual CES demand (taking into account industry weights) for each industry good given by equation (2.21). The price aggregator also takes the usual form given by equation (2.22). Note that weights \(a_j\) need to be added for each industry.

\[
Y_{j,t} = a_j \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma_j} Y_t \quad (1.44)
\]

\[
P_t = \left( \sum_{j=1}^{J} a_j P_{j,t}^{1-\sigma_j} dj \right)^{\frac{1}{1-\sigma_j}} \quad (1.45)
\]

**Industry Aggregator** I allow for different industries with different weights and degrees of price rigidity. This is done for two reasons. The primary reason is that allowing for different degrees of price rigidities increases the degree of monetary non-neutrality which is otherwise unrealistically low. See Carvalho (2006) for a detailed discussion. It is also more realistic to allow for different industries with different degrees of price rigidity.

A perfectly competitive firm of firm \(j\) aggregates all the intermediate goods in that industry to produce the good for sector \(j\). The sector firm has the following production function:

\[
Y_{j,t} = \left( \int_0^1 Y_{i,j,t}^{\frac{\sigma_j-1}{\sigma_j}} di \right)^{\frac{\sigma_j}{\sigma_j-1}}
\]

Therefore, the industry aggregator has the usual CES demand for each intermediate good given by equation (2.23). The price aggregator also takes the usual form given by equation (2.24).

\[
Y_{i,j,t} = Y_{j,t} \left( \frac{P_{i,j,t}}{P_{j,t}} \right)^{-\sigma} \quad (1.46)
\]

\[
P_{j,t} = \left( \int_0^1 P_{i,j,t}^{1-\sigma} di \right)^{\frac{1}{1-\sigma}} \quad (1.47)
\]
Intermediate Goods Firms Cost Minimisation  The output of an intermediate firm \(i\) in industry \(j\) at time \(t\) is given by equation (1.48). Intermediate firms have Cobb Douglas production over capital \((K_{i,j,t})\) and labor \((L_{i,j,t})\). Productivity is denoted \(A_t\).

\[
Y_{i,j,t} = A_t K_{i,j,t}^\alpha L_{i,j,t}^{1-\alpha} \quad (1.48)
\]

Real profits of an intermediate firm \(\Omega_{i,j,t}\) in a single period are given by equation (1.49). They rent capital from consumers at real rate \(r_t\). They also have to refund consumers for the depreciation \(\delta\) in capital. They pay workers a real wage \(W_t\) for each unit of labor. A tax (surplus) \(\tau\) on renting capital and labor is introduced. In equilibrium, the lump sum transfer is set so that each period the amount transferred to the firm equals the tax (subsidy) it paid (received) on renting capital and labor (so the only impact of the tax is to adjust the cost of production for the firm). Firms do not observe that the tax will be transferred back to them hence why the transfer is shown in curly brackets in equation (1.49).

\[
\Omega_{i,j,t} = \frac{P_{i,j,t} Y_{i,j,t}}{P_t} - (1 + \tau)((r_t + \delta)K_{i,j,t} + W_t L_{i,j,t}) + \{\tau((r_t + \delta)K_{i,j,t} + W_t L_{i,j,t})\} \quad (1.49)
\]

Intermediate firms minimise costs in the standard manner, which requires that equations (1.50) and (1.51) hold. \(MC_t\) represents the marginal cost of the firm before tax. The problem is shown in detail in appendix A.2.1

\[
MC_t = \frac{r_t + \delta}{\alpha A_t K_t^{\alpha-1} L_t^{1-\alpha}} \quad (1.50)
\]

\[
MC_t = \frac{W_t}{(1 - \alpha) A_t K_t^\alpha L_t^{-\alpha}} \quad (1.51)
\]

Output and profits (equations (1.48) and (1.49)) can be aggregated to get equations (1.52) and (2.26). It is also possible to write profits in the more usual form given in equation (1.54).

\[\text{17}\text{I introduce the tax so the equilibrium real rate can be set to take a particular value.}\]
These steps are discussed in appendix A.2.1

\[ Y_t^\nu_t = A_tK_t^{\alpha}L_t^{1-\alpha} \quad (1.52) \]

\[ \Omega_t = Y_t - Y_tMC_t^\nu_t \quad (1.53) \]

\[ \Omega_t = Y_t - (r_t + \delta)K_t - W_tL_t \quad (1.54) \]

As part of the aggregation of output and profits it is necessary to define a price dispersion variable \( \nu_t \) (defined in equation (2.27)) which in turn aggregates the price dispersion of individual industries \( \nu_{j,t} \) (defined in equation (2.28)).

\[ \nu_t = \sum_{j=1}^{J} a_j \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma_2} \nu_{j,t} dj \quad (1.55) \]

\[ \nu_{j,t} = \int_{0}^{1} \left( \frac{P_{i,j,t}}{P_{j,t}} \right)^{-\sigma} di \quad (1.56) \]

**Rewriting Cost Minimisation Conditions in terms of the Markup** The average markup \( m_t \) is defined to be the inverse of the marginal cost of producing one final good i.e. equation (1.57)\(^\text{18}\). Profits (equation (2.26)) are written in terms of the markup in equation (1.58). Using equation (1.52) as well, equations (1.50) and (1.51) can be rewritten in terms of \( m_t \) as equations (1.59) and (1.60).

\[ m_t = \frac{1}{MC_t^\nu_t} \quad (1.57) \]

\[ \Omega_t = (1 - \frac{1}{m_t})Y_t \quad (1.58) \]

\(^{18}\)This includes the degree of price dispersion because as the price dispersion increases, demand for intermediate goods with cheaper prices rises even though these goods contribute less to making a final good than less used goods with more expensive prices. Thus, more intermediate goods must be used to produce a final good than in the case where there is no price dispersion.
\[
\frac{\alpha}{m_t} = \frac{(r_t + \delta)K_t}{Y_t} \quad (1.59)
\]
\[
\frac{1 - \alpha}{m_t} = \frac{W_tL_t}{Y_t} \quad (1.60)
\]

**Intermediate Firm Profit Maximisation**  Firms in each industry \( j \) have a \( \lambda_j \) probability of updating their price each period. When they do get to change their price, firms maximise equation (2.25) subject to the demand for their good from the industry aggregator firm (equation (2.23)). Firms discount future real profits by a fixed amount \( \frac{\beta_f}{\bar{R}} \). The \( \frac{1}{\bar{R}} \) represents the risk-free discount of the future. Firms are allowed to discount by an additional \( \beta_f \). Therefore, firms maximise equation (2.25) subject to the demand from industry aggregator firms (equation (2.23)).

\[
\max_{P_{j,t}^*, \nu_{j,t}} \sum_{k=0}^{\infty} \left( \frac{\beta_f}{\bar{R}} \right)^k (1 - \lambda_j)^k \left[ \frac{P_{j,t}^* Y_{i,j,t+k}}{P_{t+k}} - (1 + \tau)MC_{j,t+k}Y_{i,j,t+k} \right] \quad (1.61)
\]

**Rewriting Price Evolution Equations**  Equation (2.28) can be rewritten as equation (2.29). Equation (2.24) can be rewritten as equation (2.30). There is a relationship between inflation in an industry and the relative price in that industry that holds by definition and is shown in equation (1.64). And equation (2.22) can be rewritten as equation (2.31). These steps are discussed in appendix A.2.1.

\[
\nu_{j,t} = \lambda_j \left( \frac{P_{j,t}^*}{P_{j,t}} \right)^{-\sigma} + (1 - \lambda_j)\nu_{j,t-1}\Pi_{j,t}^\sigma \quad (1.62)
\]

\[
1 = \lambda_j \left( \frac{P_{j,t}^*}{P_{j,t}} \right)^{1-\sigma} + (1 - \lambda_j)\Pi_{j,t}^{-\sigma-1} \quad (1.63)
\]

\[
\Pi_{j,t} = \frac{P_{j,t}}{P_t} \frac{P_{t-1}}{P_{j,t-1}} \quad (1.64)
\]

\[
1 = \sum_{j=1}^{J} \alpha_j \left( \frac{P_{j,t}}{P_t} \right)^{1-\sigma_2} \quad (1.65)
\]
**Intermediate Firm Profit Maximisation Solution**  

The solution to equation (2.25) can be rewritten as the first-order condition (equation (1.66)) plus two auxiliary equations (equations (1.67) and (1.68)). The derivation is discussed in appendix A.2.1.

\[ U_{j,t} \frac{P_{j,t}}{P_t} - V_{j,t} = 0 \]  

\[ U_{j,t} = \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma_2} Y_t + \mathbb{E}_t \left[ \frac{\beta f}{R} (1 - \lambda_j) \Pi^\sigma_{j,t+1} \Pi_t^{-1} U_{j,t+1} \right] \]  

\[ V_{j,t} = \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma_2} Y_t \frac{\sigma}{\sigma - 1} (1 + \tau) MC_t + \mathbb{E}_t \left[ \frac{\beta f}{R} (1 - \lambda_j) \Pi^\sigma_{j,t+1} V_{j,t+1} \right] \]

1.3.3 Monetary and Fiscal Policy

When investigating the long-run equilibrium, a monetary rule does not need to be specified (since we’re just computing the steady state). In this case, just note that the central bank holds inflation at some target \( \pi^* \). However, a monetary rule is needed when investigating the equilibrium with shocks. A similar monetary rule to Coibion et al. (2012) is used which is given in equation (1.69).

\[ I_t = \max \{ I_{t-1}^{\rho_1} I_{t-2}^{\rho_2} (\bar{\Pi}) \left( \frac{\Pi_t}{\Pi} \right)^{\phi_y} \left( \frac{Y_t}{Y} \right)^{\phi_y} (1 - \rho_1 - \rho_2, 0) \} \]

The government is assumed to have no debt/savings:

\[ B_t = 0 \]

1.3.4 Other Conditions

In the main model, \( A_t = 1 \).

Total labor is just the population-weighted sum of labor given by equation (1.70). In the

---

19 The one change is that the interest rate responds to the difference from output from its steady state rather than its natural level.
exogenous labor case, $L_t$ is effectively fixed.

\[ L_t = \frac{\sum_{i=0}^{M-1} \left(\frac{1}{1+n}\right)^i L_{i,t}}{\sum_{i=0}^{M-1} \left(\frac{1}{1+n}\right)^i} \]  

(1.70)

### 1.4 Model Solution and Calibration

In this section, I discuss how the conditions derived in section 1.3 can be used to do policy analysis.

#### 1.4.1 Full Conditions

In this subsection, the conditions derived in section 1.3 are summarized.

The household’s problem is summarized by $2M + 4$ conditions: $M - 1$ Euler condition(s) (equation (1.26)), two arbitrage conditions (equations (1.25) and (1.29)), the sum of savings shares (equation (1.42)), the amounts of savings and assets (equations (1.40) and (1.41)) and $M$ simplified budget constraints (equation (1.43)).

The firm’s cost minimisation problem is summarized by 4 conditions: the cost minimisation conditions (equations (1.50) and (1.51)), the definition of output (equation (1.48)) the definition of profits (equation (1.49)).

For the firm’s pricing problem: There is a condition for each industry for equations (1.64), (1.66) to (1.68), (2.29) and (2.30). There are also two overall conditions (equations (2.27) and (2.31)). In total, the firm’s pricing problem is summarized by $6J + 2$ conditions.

There is one condition from monetary policy (equation (1.69)) and one equilibrium condition (equation (1.70)).

In total, there are $2M + 6J + 12$ conditions. These correspond to the following variables:

\[
\{C_{i,t}\}_{i=0}^{M-1}, \{s_{i,t}\}_{i=1}^{M-1}, I_t, \Pi_t, R_t, W_t, MC_t, K_t, L_t, Y_t, Z_t, \Omega_t, S_t, T_t, \nu_t, \left\{ \frac{P_{j,t}}{P_t}, \frac{P^{*}_{j,t}}{P_t}, \Pi_{j,t}, \nu_{j,t}, U_{j,t}, V_{j,t} \right\}_{j=1}^{J}
\]
1.4.2 Steady State

In this subsection, the long-run equilibrium (the steady state) of the model is computed. I solve for the steady state in a similar manner to section 1.2. The relative asset demand and relative asset supply are computed and then I find equilibria where they intersect. Relative assets $a_t$ are defined to be total savings held by agents at the end of a period ($S^p_t$) divided by labor income ($W_tL_t$). This is shown in equation (1.71). The reason relative assets are used is because then asset demand doesn’t depend upon the wage which makes the model easier to solve. In graphs, references are made to ”annualized assets” which are just assets divided by annualized labor income rather than labor income for one period.

$$a_t = \frac{S^p_t}{W_tL_t} \quad (1.71)$$

The solution is broken into three parts. Firstly, the markup $\bar{m}$ is solved for given the inflation target; this is explained in appendix A.3.1. Secondly, the supply of relative assets $\bar{a}^s$ is solved for given the markup; this is explained in appendix A.3.2. Thirdly, the demand for relative assets $\bar{a}^d$ is solved for; this is explained in appendix A.3.3. It is then possible to find the steady state by looking for points where the supply and demand for relative assets intersect.

I demonstrate that the equilibrium must exist, is dynamically efficient and satisfies $\bar{R} > 1 + n$ in appendix A.4

1.4.3 Shocks and Log-linearized Conditions

In this subsection, shocks are incorporated into the log-linearised versions of the full set of model conditions found in section 1.4.1.

Similar shocks to Coibion et al. (2012) are incorporated into the model.\footnote{The differences are that this model does not have a government sector so it does not have no government shocks and productivity growth does not follow some trend here.} The shocks are
to technology, the risk premium, the Phillips Curve and the nominal interest rate and are denoted as respectively $\epsilon_{a,t}, \epsilon_{q,t}, \epsilon_{m,t}, \epsilon_{i,t}$ with standard deviations of the shocks respectively given as $\sigma_a, \sigma_q, \sigma_m, \sigma_i$. The productivity, the risk premium and the Phillips Curve shocks are AR(1) processes:

$$\hat{A}_t = \rho_a \hat{A}_{t-1} + \epsilon_{a,t}$$
$$\hat{q}_t = \rho_q \hat{q}_{t-1} + \epsilon_{q,t}$$
$$\hat{m}_t = \rho_m \hat{m}_{t-1} + \epsilon_{m,t}$$

Denote the log linearisation of $X$ around steady state as $\hat{X}$:

$$\hat{X}_t = \log(X_t) - \log(\bar{X})$$

Household conditions (equations (1.25), (1.26), (1.29) and (1.40) to (1.43)) are log-linearized to yield:

$$\mathbb{E}_t[\hat{C}_{i+1,t+1}] = \frac{1}{\gamma}(\mathbb{E}_t[\hat{R}_{t+1}] + \hat{q}_t) + \hat{C}_{i,t}$$
$$\mathbb{E}_t[\hat{R}_{t+1}] = \hat{I}_t - \mathbb{E}_t[\hat{\Pi}_{t+1}]$$

$$\hat{Z}_t + \mathbb{E}_t[\hat{R}_{t+1}] = \frac{\bar{\Omega}}{\bar{\Omega} + \bar{Z}} \mathbb{E}_t[\hat{\Pi}_{t+1}] + \frac{\bar{Z}}{\bar{\Omega} + \bar{Z}} \mathbb{E}_t[\hat{Z}_{t+1}]$$

$$0 = \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \bar{s}_i \hat{s}_{i,t}$$

$$\bar{S}^p \hat{s}^p_t = \frac{\bar{B}}{\bar{P}} \left( \frac{B_{t+1}}{P_{t-1}} \right) + \bar{K} \hat{K}_{t+1} + \bar{Z} \hat{Z}_t$$

$$\hat{C}_t \hat{C}_{i,t} + s^p_{i+1,t} \hat{s}^p_t\hat{s}^p_{i+1,t+1} = \bar{W} \bar{L}_t \hat{W}_t + s^p_t \bar{T} \bar{s}^p_{i+1,t-1} + \hat{T}_t$$

Firm cost minimization conditions (equations (1.50) to (1.52) and (2.26)) are log-linearized
to yield:

\[ \dot{MC}_t = \dot{r}_t - \dot{A}_t + (1 - \alpha)\dot{K}_t \]

\[ \dot{M}C_t = \dot{W}_t - \dot{A}_t - \alpha\dot{K}_t \]

\[ \dot{Y}_t + \dot{\nu}_t = \dot{A}_t + \alpha\dot{K}_t \]

\[ \dot{\Omega}_t = \dot{Y}_t - \frac{\dot{MC} \dot{\nu}}{1 - \dot{MC} \dot{\nu}} (\dot{MC}_t + \dot{\nu}_t) \]

Profit maximization conditions (equations (1.64), (1.66) to (1.68), (2.27) and (2.29) to (2.31)) are log-linearized to yield:

\[ \bar{\nu}_j \hat{\nu}_{j,t} = -\sigma \lambda_j \left( \frac{P^*_j}{P_j} \right)^{-\sigma} \left( \frac{P^*_j}{P_j} \right)^{1-\sigma} (\dot{P}_{j,t} P_{j,t}^*) + (1 - \lambda_j)\bar{\nu}_j \bar{\Pi}^{\sigma} (\hat{\nu}_{j,t-1} + \sigma \bar{\Pi}_{j,t}) \]

\[ 0 = (1 - \sigma) \bar{\lambda}_j \left( \frac{P^*_j}{P_j} \right)^{1-\sigma} \left( \frac{P^*_j}{P_j} \right)^{-\sigma} + (\sigma - 1)(1 - \lambda_j)\bar{\Pi}^{\sigma-1} \hat{\Pi}_{j,t} \]

\[ \bar{\Pi}_{j,t} = \left( \frac{P_{j,t}}{P_t} \right) + \hat{\Pi}_t - \left( \frac{P_{j,t-1}}{P_{t-1}} \right) \]

\[ \bar{U}_{j,t} + \left( \frac{P^*_j}{P_j} \right) + \left( \frac{P^*_j}{P_j} \right) = \hat{V}_{j,t} \]

\[ \bar{U}_{j,t} \hat{U}_{j,t} = \left( \frac{P_j}{P} \right)^{-\sigma_2} \bar{Y} \left( -\sigma_2 \frac{P_{j,t}}{P_t} + \hat{Y}_t \right) + \hat{\beta}_{j} \left( 1 - \lambda_j \right) \bar{\Pi}^{\sigma-1} \bar{U} \left( \sigma \bar{\Pi}_{j,t+1} - \hat{\Pi}_{t+1} + \hat{U}_{j,t+1} \right) \]

\[ \bar{V}_{j,t} \hat{V}_{j,t} = \left( \frac{P_j}{P} \right)^{-\sigma_2} \bar{Y} \frac{\sigma}{\sigma - 1} (1 + \tau) \dot{MC} \left( -\sigma_2 \frac{P_{j,t}}{P_t} + \hat{Y}_t + \dot{MC}_t + \hat{m}_t \right) + \hat{\beta}_{j} \left( 1 - \lambda_j \right) \bar{\Pi}^{\sigma} \bar{V} \left( \sigma \bar{\Pi}_{j,t+1} + \hat{V}_{j,t+1} \right) \]

\[ 0 = \sum_{j=1}^{J} (1 - \sigma_2) \alpha_j \left( \frac{P_j}{P} \right)^{1-\sigma_2} \left( \frac{P_{j,t}}{P_t} \right) \]

\[ \bar{\nu} \hat{\nu}_t = -\sum_{j=1}^{J} \sigma_2 \alpha_j \left( \frac{P_t}{P} \right)^{-\sigma_2} \bar{\nu}_j \left( \frac{P_{j,t}}{P_t} \right) + \hat{\nu}_{j,t} \]

34
The monetary rule (equation (1.69)) is log-linearized:

\[ \hat{I}_t = \max \{ \rho_1 \hat{I}_{t-1} + \rho_2 \hat{I}_{t-2} + (1 - \rho_1 - \rho_2)(\phi_x \hat{I}_t + \phi_y \hat{Y}_t) + \epsilon_{i,t}, -\log(I) \} \] (1.72)

Labor in equilibrium (equation (1.70)) is log-linearized:

\[ \hat{L}_t = \frac{\sum_{i=0}^{M-1} (\frac{1}{1+n})^i \hat{L}_{i,t}}{\sum_{i=0}^{M-1} (\frac{1}{1+n})^i} \]

1.4.4 Policy Functions and Simulation

In this subsection, the method used to simulate the log-linearized model is discussed.

Firstly, note that a standard DSGE model has been derived. Therefore, a first-order linear perturbation of the conditions given in section 1.4.3 can be applied. However, any first-order perturbation will ignore the ZLB in equation (1.72).

To capture the impact of the ZLB, a similar method to Guerrieri and Iacoviello (2015) is used. The basic idea is that agent’s choices at time \( t \) are solved for by finding what they would do without shocks in the future if there was no ZLB. If the central bank would set a nominal interest rate below zero in this case then the estimation is rerun with the central bank constrained in certain periods to be at the ZLB. This process continues until a situation where the central bank is not setting a nominal interest rate below zero is found. The exact algorithm used is as follows:

1. Set \( t \) to be the first period of the simulation.
2. Assume that at \( t + 100 \), without any further shocks, the ZLB will no longer bind.
3. Guess the following regime: The ZLB does not bind in any period from \( t \) to \( t + 99 \).
4. Solve backwards from \( t + 99 \) to \( t \) to get the policy functions for each period under the guessed regime.\(^{21}\)

\(^{21}\)The computations used to solve backwards from \( t + 99 \) to \( t \) is the same as in Guerrieri and Iacoviello.
5. Using the policy functions computed in step 3, solve out for the path of the economy from $t$ to $t + 99$ in the case where there are no shocks from $t + 1$ to $t + 99$.

6. Verify whether the nominal interest rate is always greater than or equal to zero in every period. If it is not:

   (a) If this regime is such that the economy was never at the ZLB, set that the regime is now such that the ZLB binds at $t$ but not in future periods. If this regime is such that the economy was at the ZLB until period $t + s$, set that the regime is now such that the economy is at the ZLB until $t + s + 1$.

   (b) Go back to step 3.

7. Take agent’s choices at $t$ to be the simulation values for $t$. Now set $t$ to be the next period of the simulation and go back to step 2.

The difference between my method and Guerrieri and Iacoviello (2015) is in step 6. In my method, step 6 implies that if a central bank knows that it would hit the ZLB in the future without additional shocks, it will lower its nominal interest rate to zero in all preceding periods. Guerrieri and Iacoviello’s algorithm implies that the ZLB should only bind in periods when the nominal interest rate would have been below zero according to the monetary rule. This means the ZLB can bind in nonconsecutive periods and can bind in the future even if it doesn’t bind today.

Guerrieri and Iacoviello warn that their method is a simple first pass and I find there are advantages to adapting their method to what I have described here. Firstly, since it is more general, the Guerrieri and Iacoviello method often does not converge in my model. Secondly, my method works more quickly. Thirdly, I think it is more realistic to assume that

---

22I have not encountered the situation where the ZLB binds from $t$ to $t + 99$.

23My algorithm could also be considered as a simpler version of Andrade et al. (2018) who also adapt the algorithm of Guerrieri and Iacoviello (2015). They allow for the ZLB to start binding at some period $t_1 \geq t$ and then stop binding at some later period $t_2 > t_1$. I effectively set that $t_1 = t$ always.
if a central bank knows that the economy is likely to hit the ZLB in the future, the central bank will lower nominal interest rates to zero immediately rather than waiting.

1.4.5 Calibration

**General Parameters** Each period is set to represent a quarter. Standard parameters are set as follows: \( \alpha = 0.3, \beta = 0.98^{\frac{1}{4}}, \delta = 0.1^{\frac{1}{4}}, \gamma = 2. \) \( \xi \) is set to be 0 so there is no premium on safe bonds in the baseline calibration.

\( M = 220 \) is calibrated to capture each quarter of life of an adult between the ages of 24 and 78. The simulation begins at age 24 to avoid having to worry about how to capture college. Agents’ last year of life is 78 because the life expectancy of someone in the US is currently just under 79 years.

With exogenous labor supply, \( \bar{L}_i \) (hours worked by each age) is set to match the average hours worked of a person of that age in the American Time Use Survey between 2003 and 2016. With endogenous labor supply: \( x_i \) in the disutility of labor function (equation (1.23)) is set so that when \( \beta \bar{R} = 1 \) we have that \( \bar{L}_i \) matches the hours worked in the exogenous case. This is explained in appendix A.5.1. \( \eta \) is set to be 1 in the endogenous labor supply case.

The industry weights and frequencies of price adjustment are set to match regular non-sale prices in \cite{Nakamura and Steinsson 2008}. The elasticity of substitution between varieties within industries \( (\sigma) \) is set to be \( \sigma = 8 \). This is in between the lower and upper bounds of respective 6 and 10 used in \cite{Carvalho et al. 2016}. The elasticity of substitution between industries \( (\sigma_2) \) is set to 1 as in \cite{Shamloo 2010}.

It is important that firm discounting is set correctly since it makes a difference for the size of the first part of the channel. The degree of firm discounting is based upon the Weighted Average Cost of Capital (WACC) which is the average a company is expected to pay to finance its assets. \cite{Jagannathan et al. 2016} estimated that it was 8% in 2003 when the

\[ ^{24} \text{I actually set it to be 1.001 otherwise I would have to rewrite the indices since 1 is a special limiting case.} \]

\[ ^{25} \text{This is used since it is the cost to the firm of not obtaining funds earlier by setting a lower markup.} \]
expected ten year rate on real bonds \( (r^e) \) was 2.8p.p. Graham and Harvey (2011) estimated it was 10.0\% in 2011Q1 when \( r^e \) was 2.2\%. Graham and Harvey (2012) then estimated it was 9.3\% in 2012Q2 when \( r^e = 1.3\% \). From these three surveys, the average wedge between WACC and the expected real rate is 7p.p. Therefore, a firm discount of \( \beta_f = \left( \frac{1}{1.07} \right)^{\frac{1}{4}} \) is applied.

\( \bar{r} \) is set to be 2.06\% when \( \pi^* = 2\% \). This matches the average real interest rate on treasury bills between 1995 and 2007. \( \tau \) (the tax on the labor and capital inputs) is calibrated to set \( \bar{r} \) at this level.

**Simulation Parameters** The parameters needed to solve for the long-run (steady state) equilibrium have been fully described. I now describe the parameters that are only needed for the simulation.

The same monetary rule parameters are used as in Coibion et al. (2012): \( \rho_{i,1} = 1.05, \rho_{i,2} = -0.13, \phi_\pi = 2.5, \phi_y = 0.11 \).

Where possible, the shock parameters are set to be the same as in Coibion et al. (2012). Therefore: \( \sigma_a = 0.009, \sigma_q = 0.0024, \sigma_m = 0.0014, \sigma_i = 0.0024, \rho_q = 0.947, \rho_m = 0.9 \). The persistence in productivity is set to be \( \rho_a = 0.9 \).

### 1.5 Impact of Raising the Inflation Target

In this section, I consider the impact of raising the annual inflation target from 2\% to 4\%. 2\% is chosen to be the baseline level of the inflation target because that is the standard inflation target among developed countries. The impact of raising the inflation target by 2p.p. to 4\% is investigated since that is the most commonly proposed adjustment (Blanchard et al. (2010), Ball (2014), Krugman (2014)).

Figure 1.5 shows the impact of the policy experiment on the supply and demand for relative assets. The impact of raising the inflation target has exactly the same qualitative impact as in section 1.2. The supply of assets shifts left since a lower markup lowers profits
and thus the value of firms. Therefore, there are fewer assets available for households to hold. It does not shift the demand for assets by households. Observe that a shift left in the supply of savings lowers the equilibrium real rate and equilibrium relative assets. The intuition for the fall in the equilibrium real rate is also the same as in the simplified model. Households rely upon savings to consume when they are old. A fall in savings means that households consume relatively less when they are old so the price of saving rises which is equivalent to a fall in the equilibrium real rate.

Figure 1.6 shows the impact of the rise in inflation on the consumption path of agents across their lives. A rise in the inflation target lowers the consumption of the old relative to the young. This is because agents save less for when they are old as a result of the lower supply of assets.

Table 1.1 shows the numerical impact of the policy experiment with the default calibra-
A rise in the inflation target leads to a fall in the markup of 1.07p.p. This is just the first part of the channel where when inflation increases firms set a lower markup due to price rigidities. Ceteris paribus, a fall in savings reduces the ability of older agents to consume. Therefore, the consumption of the old falls relative to the young. The second row of table 1.1 shows that the consumption of the older half of consumers falls by 5.14 percent relative to the younger half of consumers. Next, observe that agents hold 4.46 percent more capital (relative to the representative agent baseline case). Agents want to save more in capital to try to reduce their loss of consumption when they are old. Since agents want to save more, the price of saving (the equilibrium real rate) falls. In this case, it falls by 0.38p.p.

Raising the inflation target would be less effective in reducing the probability of the ZLB. A rise in the inflation target of 2p.p. only raises the equilibrium nominal interest rate.

---

26 By row, the mathematical expressions for what table 1.1 shows are: $100\Delta m$, $100\Delta \log(\sum_{i=120}^{239} C_i) - \log(\sum_{i=119}^{119} C_i)$, $100(\Delta \log(K_{old}) - \log(K_{old}))$, $100\Delta r$. The change in capital is relative to the case when $\bar{r}$ does not change (which includes the representative case) since I want to demonstrate that agents choose to hold more capital than when $\bar{r}$ does not change.
Table 1.1: Policy Experiment with Default Calibration

<table>
<thead>
<tr>
<th>Defaults</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in markup (p.p.)</td>
<td>-1.07</td>
</tr>
<tr>
<td>Change in $\frac{C_{old}}{C_{young}}$ (%)</td>
<td>-5.14</td>
</tr>
<tr>
<td>Change in r (p.p.)</td>
<td>-0.38</td>
</tr>
<tr>
<td>Change in K (%)</td>
<td>4.46</td>
</tr>
</tbody>
</table>

by 1.62p.p. in the default calibration, as opposed to the 2p.p. rise widely assumed and predicted by standard models. Since nominal interest rates would rise by less, this would give policymakers less room to cut in bad times before hitting the ZLB.

Figure 1.7 shows how lowering the intertemporal elasticity of substitution (IES) to 0.1 affects the supply and demand of relative assets. Recent estimates from [Best et al. (2018)] suggest that the IES is 0.1. Lowering the IES to 0.1 (from its baseline value of 0.5) causes the demand for assets to tilt backwards. The reason for this is because when agents have low IES, they have a stronger desire to consume the same each period. When the real interest rate rises, agents get a higher return on their savings allowing them to consume more when they are old. With a low IES, they will then reduce the amount they save to rebalance consumption back to when they are young. In this sense, the income effect of raising the real rate dominates when IES is high enough.

Table 1.2 shows the numerical impact of the policy experiment under different IES. Note that the column with an IES of 0.5 is just the default calibration and matches the results in table 1.1. A lower IES which pushes agents to consume the same in each period implies that, when consumption of the old falls, the price of savings rises by more and thus the return on savings falls by more. With an IES of 0.1, the equilibrium real rate falls by 0.67p.p. compared to 0.38p.p. in the baseline case. Agents can mitigate the fall in their consumption when they are old by raising their investment in capital. Since agents with low IES really want to consume the same over time, they invest more in capital hence why it

---

27The rows have the same mathematical expressions as footnote 26.
rises by 8.02% relative to the representative agent case compared to 4.46% in the baseline calibration.\textsuperscript{28}

Table 1.3 explores the impact of allowing for endogenous labor supply.\textsuperscript{29} Allowing for

Table 1.2: Impact of Changing Intertemporal Elasticity of Substitution on Results of Policy Experiment

<table>
<thead>
<tr>
<th>IES ($\frac{1}{4}$)</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.5</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in markup (p.p.)</td>
<td>-1.04</td>
<td>-1.05</td>
<td>-1.06</td>
<td>-1.07</td>
<td>-1.09</td>
</tr>
<tr>
<td>Change in $\frac{C_{old}}{C_{young}}$ (%)</td>
<td>-1.83</td>
<td>-3.11</td>
<td>-4.01</td>
<td>-5.14</td>
<td>-6.35</td>
</tr>
<tr>
<td>Change in K (%)</td>
<td>8.02</td>
<td>6.80</td>
<td>5.84</td>
<td>4.46</td>
<td>2.74</td>
</tr>
<tr>
<td>Change in r (p.p.)</td>
<td>-0.67</td>
<td>-0.58</td>
<td>-0.50</td>
<td>-0.38</td>
<td>-0.24</td>
</tr>
</tbody>
</table>

\textsuperscript{28}The markup falls by less when IES is lower. This is because when IES is lower, the real interest rate falls by more which implies discounting falls and thus there’s the impact of inflation on the markup is lessened slightly.

\textsuperscript{29}Change in $\frac{C_{old}}{C_{young}}$ has the following mathematical expression: $100 \Delta \pi \left( \log(\sum_{i=100}^{239} L_i) - \log(\sum_{i=100}^{240} L_i) \right)$. The other rows have the same mathematical expressions as footnote 26.
Table 1.3: Impact of Changing Elasticity of Labor Supply on Results of Policy Experiment

<table>
<thead>
<tr>
<th>Elasticity of Labor Supply ($\eta$)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in $\frac{L_{old}}{L_{young}}$ (%)</td>
<td>-2.85</td>
<td>-3.68</td>
<td>-4.07</td>
<td>-4.45</td>
<td>-4.77</td>
</tr>
<tr>
<td>Change in $\frac{L_{old}}{L_{young}}$ (%)</td>
<td>4.65</td>
<td>3.01</td>
<td>2.22</td>
<td>1.45</td>
<td>0.78</td>
</tr>
<tr>
<td>Change in K (%)</td>
<td>2.17</td>
<td>2.91</td>
<td>3.28</td>
<td>3.67</td>
<td>4.02</td>
</tr>
<tr>
<td>Change in r (p.p.)</td>
<td>-0.21</td>
<td>-0.27</td>
<td>-0.30</td>
<td>-0.33</td>
<td>-0.35</td>
</tr>
</tbody>
</table>

Endogenous labor leads to a somewhat smaller fall in the real interest rate than in the baseline calibration with exogenous labor (figure 1.5). The reason for this is that when the markup falls and thus savings falls so that agents consume less when they are old relative to when they are young, agents can choose to work more when they are old to substitute for the loss in consumption when they are old. The extent to which they do this depends upon the elasticity of labor supply. With a low elasticity of labor supply, they work 4.65% relatively more when they are old than when they are young. Consequently, with a low elasticity of labor supply, households also choose to invest relatively less in capital since they are replacing consumption when they are old through working when they are older instead. The fall in the equilibrium real rate lessens when households have a low elasticity of labor supply. As the elasticity of labor supply gets large, the model converges back to the exogenous labor supply case in table 1.1.

I also explore the impact of idiosyncratic labor shocks. Appendix A.7 examines how idiosyncratic labor shocks can be embedded within the New Keynesian life cycle model of section 1.3. Figure 1.8 shows the impact of adding idiosyncratic labor shocks. The blue, orange and green curves are identical to figure 1.5. The blue and orange curves are the supply of relative assets respectively before and after the shift left in relative asset supply due to a rise in the inflation target. The green curve is the demand for relative assets with OLG households and no idiosyncratic labor shocks. The dashed red curve represents the

---

30 Define old agents to be those above the average age and those who are young to be below the average age.
OLG model with idiosyncratic labor included. The only effective impact is that the demand for relative assets shifts out. The intuition for this is that households face more risk so they want to save more as a precaution against this risk. However, there does not appear to be a substantive impact on the degree to which the real interest rate falls following the shift left in the supply of assets.

### 1.6 Optimal Inflation Target

In this section, I analyse the optimal inflation target. The main aim here is to investigate how the optimal inflation target changes in response to a fall in the equilibrium real interest rate. I am interested in this because much of the motivation for raising the inflation target has focused upon the apparent recent fall in the equilibrium real rate in much of the developed world which makes the probability of hitting the ZLB more likely. I also investigate the level of the optimal inflation target.
To consider these questions, the economy is simulated over 1,000 periods (250 years). A set of shocks is drawn from the calibrated shock distributions and these same shocks are used in every simulation. The welfare (the population-weighted utility averaged across all periods) is then computed under different inflation targets. To capture the impact of a fall in the equilibrium real rate, we investigate what happens when agents' premium on safe bonds rises by 2p.p. which implies that $\xi = -0.02$. I choose to model a fall in the equilibrium real rate as an increase in the desire for safe assets because this is often used to explain the fall in the equilibrium real rate and because this does not have any impact upon asset demand or asset supply so it is easier to isolate the impact of the change.

The welfare under the baseline calibration is displayed in figure 1.9. The blue line represents the welfare under the baseline equilibrium real rate (2.07%). The orange line represents the welfare under a 2p.p. lower equilibrium real rate. For a low inflation target, the welfare under the lower equilibrium real rate is lower. This is because under low levels of inflation a lower equilibrium real rate implies a higher probability of hitting the ZLB which implies lower welfare. However, for a higher inflation target, the welfare is identical. Figure 1.10 shows the probability of hitting the ZLB in each case. The probability is higher for the lower equilibrium real rate case when the inflation target is low. However, once the inflation target is high enough, there is no probability of hitting the ZLB in either case so the welfare is identical. The fall in the equilibrium real rate implies that the optimal inflation target rises by 0.3p.p. from 0.6% to 0.9%.

The baseline calibration does not generate shocks that hit the ZLB very frequently. Larger shocks are allowed for by scaling up each of the shock standard deviations by a factor of 1.5. The revised welfare and probability of hitting the ZLB are shown in figure 1.11 and figure 1.12 respectively. The larger shocks mean that under a low equilibrium real rate, there is a higher probability of hitting the ZLB. This implies that the optimal inflation target rises by more when the equilibrium real rate falls. The fall in the equilibrium real rate of 2p.p. leads to a rise of 0.6p.p. in the optimal inflation target from 0.6p.p. to 1.2p.p.
Figure 1.9: Welfare under Baseline Calibration

Figure 1.10: Probability of Hitting the ZLB under Baseline Calibration
Figure 1.11: Welfare under Larger Shocks

![Graph showing Welfare under Larger Shocks]

Figure 1.12: Probability of Hitting the ZLB under Larger Shocks

![Graph showing Probability of Hitting the ZLB under Larger Shocks]
The simulations imply that the fall in the equilibrium real rate does not generate a large increase in the optimal inflation target. In the high shock case, the increase was only 0.6p.p. The intuition for why is that the probability of hitting the zero lower bound falls quickly when the inflation target is raised. Thus, the benefits of raising the inflation target are quickly outweighed by the increased costs of price dispersion.

In all cases that have been considered, the inflation target is not much more than 1%. This similar to the representative agent case. The optimal inflation target is low because of the high costs of price dispersion. Even once size of the shocks is increased and the ZLB binds more, the costs of inflation through higher price dispersion seem to dominate the benefits of avoiding the ZLB. This also appears to be true in the representative case since Coibion et al. (2012) compute an optimal inflation target of around 1% in a representative agent model with the ZLB.

1.7 Empirical Evidence

In this section, I provide reduced form empirical estimates of the impact of a change in trend inflation on the equilibrium real interest rate. This complements existing structural analysis. I first consider the relationship between long-run inflation and the equilibrium real rate. The mechanism I propose implies that a rise in long-run inflation lowers the equilibrium real rate. There is existing empirical evidence for this. Both King and Watson (1997) and Rapach (2003) find such a relationship using structural VAR methods. To complement this existing evidence, I conduct a reduced form analysis of the relationship.

A problem with studying the reduced form relationship between long-run inflation and the equilibrium real rate is correlated trends. Inflation has trended down in recent years at the same time as the equilibrium real rate has fallen. If the fall in inflation was the only reason for the fall in the equilibrium real rate then this would imply my channel is incorrect. However, there are many other reasons why the equilibrium real rate has fallen. Therefore,
since real rates have fallen at the same time as inflation has fallen but for reasons other than the fall in inflation, a simple regression of the equilibrium real rate on inflation is likely to produce a positively biased coefficient.

To overcome common trends in inflation and real rates, I conduct panel data regressions with time fixed effects. Using time fixed effects means that the common global trend in real rates can be controlled for. It is then possible to assess whether higher relative inflation is associated with a positive or negative deviation from the global trend in real rates. If other factors that cause deviations from the global trend in real rates for a country are uncorrelated with that country’s inflation level then this relationship is causal.

Regression: Equation (1.73) shows the estimated model. $\alpha_i$ represents country fixed effects i.e. whether the real interest rate is systematically higher in a country. $\delta_t$ represents the time fixed effects. $\beta$ is the coefficient of interest which represents the change in the real interest rate relative to the global trend associated with a 1p.p. rise in long-run inflation. Controls are also included.

$$r_{i,t} = \alpha_i + \delta_t + \beta \overline{\text{Inflation}}_{i,t} + \Gamma \text{Controls}_t + u_{i,t}$$ (1.73)

The panel is limited to OECD members. Annual data is used. Long-run inflation ($\overline{\text{Inflation}}_{i,t}$) is measured as the moving average of the current and previous four years of CPI inflation\(^{31}\)\(^{32}\) The real interest rate is measured by the 10 year expected return on safe bonds. The 10 year real rate is used since there is more data availability and its likely to be a much less noisy measure of the equilibrium real rate. To measure the 10 year real rate the measure of long-run inflation is subtracted from the nominal interest rate on 10 year

\(^{31}\)It may seem strange that the regressor is not the inflation target but nearly all inflation targets have not changed since they were introduced so the inflation target would be almost completely captured by the country fixed effects ($\alpha_i$).

\(^{32}\)Varying the measure to a different moving average does not appear to impact the results.
Table 1.4: Empirical Estimates of Relationship between Long-Run Inflation and the Equilibrium Real Rate

<table>
<thead>
<tr>
<th>RealRate_{i,t}^{10yr}</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation_{i,t-4,t}</td>
<td>-0.167***</td>
<td>-0.196***</td>
<td>-0.607***</td>
<td>-0.904***</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.044)</td>
<td>(0.069)</td>
<td>(0.149)</td>
</tr>
</tbody>
</table>

country dummies * * *
year dummies * *
econ controls *

N 1151 1151 1151 833

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

government bonds. I also allow for business cycle controls. The business cycle controls are set to be GDP growth and change in unemployment at $t$ and $t-1$.

The results are given in table 1.4. Without fixed or time effects a 1p.p. rise in long-run inflation is associated with a fall of $-0.17$p.p. in the real rate. This falls slightly once country fixed effects are introduced. It has already been noted that inflation and the real rate both have a negative trend so it is unsurprising that once time fixed effects which remove this source of positive association are added, the coefficient drops a lot to $-0.61$p.p. Controls increase the strength of the impact. A causal interpretation of the regression without controls is that a rise of 1p.p. in long-run inflation lowers the equilibrium real rate by 0.61p.p.

I verify these two relationships are robust. Table A.2 verifies the relationships continue to hold with just OECD members that joined before 1975 (these regressions exclude a number of mostly Eastern European countries that joined from the 1990s onwards). Table A.3 verifies the relationships still apply under low inflation. Table A.4 looks at whether the result remain for the period before 2000 only. Table A.5 analyzes whether the relationship

---

33Computing the measure of 10 year real interest rates by subtracting current inflation (rather than the measure of long-run inflation) from the nominal interest rate on 10 year government bonds yields similar results.

34It is undesirable to have controls that capture the long-run state of the economy since these could interfere in the long-run relationship between inflation and the real rate. Business cycle controls are short-term and should not generate this problem.
continues to hold during/after 2000 only. Spurious regressions are generally considered to
be less of a problem in panel data since we can control for common trends across countries
by time fixed effects and idiosyncratic trends within countries are unlikely to drive results.
However, I verify in table A.6 that the results still hold after differencing.

1.8 Conclusion

Many economists have proposed raising the inflation target in recent years. They argue that
this will raise the average nominal interest rate and thus reduce the probability of hitting
the ZLB. It is generally assumed and a feature of standard models that raising the inflation
target has no impact upon the equilibrium real rate.

In this paper, I show that once heterogeneity is introduced into a standard New Keynesian
model through either generational features or idiosyncratic shocks, raising the inflation target
will lower the equilibrium real rate. In my baseline calibration of a New Keynesian model
with life cycle features, a rise in the inflation target from 2% to 4% lowers the equilibrium
real rate by 0.38p.p.. Many of the arguments for raising the inflation target are premised
upon the perceived fall in the equilibrium real rate in recent years. I find that a fall of
2p.p. in the equilibrium real rate within my framework, raises the optimal inflation target
by 0.3 – 0.6p.p.

The channel I propose is empirically realistic. It relies upon the existence of price stick-
iness (or other forms of price rigidity) and the existence of a consumer life cycle. Both of
these features are observable in the real-world. There is already structural econometric ev-
idence for my channel. I provide additional reduced form evidence which further supports
its existence.

The results of this paper provide valuable insights for policy-makers. A rise in the
inflation target will lower the equilibrium real rate and therefore lead to a smaller increase
in average nominal interest rates than is generally believed. This implies that raising the
inflation target is likely to be less effective in reducing the probability of hitting the ZLB than expected. And my welfare simulations imply that a fall in the equilibrium real rate is unlikely to justify a large increase in the inflation target.

Going forward, my results imply that central banks should look at alternatives to raising the inflation target. Frameworks exist that could help to boost inflation expectations during recessions without raising long-run inflation. Price level targeting, either just for a short-term period post-recession or for all periods, appears to work theoretically. Another alternative is nominal GDP targeting, although this would have to be updated periodically to account for changes in the trend of real GDP growth. However, although these targeting frameworks could raise expectations of inflation during a recession, they do not solve the problem that central banks will likely enter future recessions with low nominal interest rates. How central banks can best tackle a slump with their main policy tool dramatically constrained remains a question that deserves more investigation.
Chapter 2

Falling Inflation and Rising Profits
2.1 Introduction

Many economists and policymakers are concerned about an apparent rise in firm profitability which has seemingly lead to lower wages and greater inequality. Common explanations for this trend are greater tolerance by competition authorities for mergers and technological changes that have produced increasing returns to scale that help monopolies to develop.

In recent years, inflation has fallen. There is an established theoretical link between inflation and profitability under the Calvo price rigidity i.e. if firms can only change their price with exogenous probability. In this paper, I explore the theoretical and empirical evidence for the negative relationship between inflation and profitability and the degree to which it can explain recent trends. I have three main results: 1. The relationship holds and is potentially stronger under non-Calvo price rigidities. 2. I explore the determinants of the size of the relationship. 3. I find estimates for how much the fall in inflation has lowered firm profitability.

Within the Calvo model, a fall in inflation raises profitability. The Calvo model implies firms can only change their price with exogenous probability. When firms do get to change their price, they effectively set a markup - their price over their marginal cost. With higher inflation, their markup will fall more quickly if they do not get to change their price since their marginal costs rise more quickly. Therefore, to maintain the same average markup, firms would need to set a higher markup when they change their price. However, firms discount the future and therefore are reluctant to set a high markup when they change their price. Thus, on average, a rise in inflation, means that firms set lower markups and thus make lower profits.

I first explore the relationship between inflation and profitability under other price rigidities than Calvo. I demonstrate that the relationship can actually stronger under menu costs. One potential reason why is that firms do not want to raise their price due to the potential of receiving idiosyncratic shocks in the future which would make a lower price more desirable.
In this case, it may be better for firms to wait to raise their price, implying lower average markups and profitability. A second potential reason why is that in the menu cost model, unlike in Calvo, the frequency of price adjustment changes for different levels of inflation. With low inflation, we can actually have a situation where firms update less frequently under menu costs than in Calvo. Since firms update less frequently, they set their prices for longer in which case the average markup could fall by more due to discounting under menu costs than Calvo. I am still investigating the reasons for why the negative relationship between inflation and the profit share can be stronger under menu costs.

Secondly, I investigate the key determinants of the size of the relationship between inflation and the profit share. Key factors are the degree of firm discounting, the degree of monetary non-neutrality and the elasticity of substitution between goods. When firms discount the future more, the negative relationship between inflation and profitability is stronger since they are then more reluctant to raise their markup when inflation rises since they care relatively more about the present. A rise in elasticity of substitution strengthens the relationship under Calvo but weakens it under Rotemberg or menu costs.

We can break recent post-Great Inflation trends in US inflation into three eras. Following the Great Inflation, there was an approximate 10 year period from 1983-1992 when inflation was controlled but still elevated. From around 1993 until the end of the Great Moderation in 2007, there was a 15 year period when inflation was approximately on target. Following the start of the Great Recession in 2008 and during the recovery from it until 2017, there has been a 10 year period where inflation has often been below target. I compare these inflation and the profit share during these eras in table 2.1.

My third main contribution is to explore how much the fall in inflation from 3.82% in 1983-1992 to 1.69% in 2008-2017 can explain the rise in the profit share from 5.86% to 9.26%. In an extended model with firm heterogeneity, a fall in inflation implies a fall of 0.47p.p. in the profit share which is 14% of the observed fall in the profit share. If we additionally allow for a more realistic degree of firm discounting raises this to 1.00p.p. or 29% of the observed
Table 2.1: Comparison of Recent Eras of Inflation

<table>
<thead>
<tr>
<th>Years</th>
<th>Average CPI Inflation</th>
<th>Average Profit Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983-1992</td>
<td>3.82</td>
<td>5.86</td>
</tr>
<tr>
<td>1993-2007</td>
<td>2.63</td>
<td>7.70</td>
</tr>
<tr>
<td>2008-2017</td>
<td>1.69</td>
<td>9.26</td>
</tr>
</tbody>
</table>

Profit share is measured by corporate profits before tax (without IVA and CCAdj). Inflation is CPI inflation. These variables are also plotted in figure B.1.

fall in the profit share.

I also demonstrate that there exists a negative empirical relationship between inflation and firm profitability through panel data regressions of the markup on trend inflation. I study this question by looking at panel data regressions of a proxy of the markup on long-run inflation controlling for country and time fixed effects in OECD countries. I find a 1p.p. rise in long-run inflation lowers the long-run markup by 0.28p.p.

There has been discussion of the theoretical link between inflation and competition. For example, King and Wolman (1996) look at the relationship between inflation and the markup under Calvo and Ascari and Sbordone (2014) investigate how a rise in trend inflation affects the marginal cost in the baseline model. I believe I am the first to investigate how other price rigidities affect the relationship, summarize the determinants of the size of the relationship and to try to obtain comprehensive model-based estimates of the effect of raising inflation on firm profitability.

Other papers have also suggested that raising the long-run inflation rate lowers the markup. Bénabou (1992) finds that raising inflation by 1p.p. lowers the markup by 0.36p.p. using a relatively reduced form approach with just US data. Banerjee and Russell (2001) apply a structural VAR approach to the G7 countries and Australia. They find that a 1p.p. rise in annual steady state inflation generates a fall of between 0.3p.p. and 2p.p. in the long-run markup.

In section 2, I introduce a simple model. In section 3, I use this simple model to examine the determinants of the size of the theoretical relationship between inflation and profitability.
In sections 4 and 5, I respectively introduce a fuller model and examine the size relationship in this case and how much the relationship can explain recent trends. Section 6 looks at empirical analysis for the relationship and section 7 concludes.

### 2.2 Simple Model

We introduce a standard simple framework to consider the determinants of the size of the relationship between inflation and profitability. We consider three different forms of price rigidity: Calvo, Rotemberg and menu costs. To keep things simple, we consider the case where there are no aggregate shocks so inflation grows at some constant rate each period.

#### 2.2.1 Single Sector Basic Setup

We use the same basic setup to investigate the implications of each form of price rigidity.

**Final Goods Firm** We have a competitive aggregator firm which produces final goods from a continuum of intermediate goods firms. It has the following problem:

\[
\min_{Y_{i,t}} \int_0^1 P_{i,t} Y_{i,t} di
\]

s.t.

\[
Y_t = \left( \int_0^1 \frac{Y_{i,t}^{1/\sigma}}{P_{i,t}} \right)^{\sigma/(\sigma-1)}
\]

We get the following first order conditions:

\[
Y_{i,t} = Y_t \left( \frac{P_{i,t}}{P_t} \right)^{-\sigma}
\]

(2.1)

\[
P_t = \left( \int_0^1 P_{i,t}^{1-\sigma} \right)^{1/\sigma}
\]

(2.2)
We can also rewrite equation (C.2) in terms of relative prices $p_{i,t} = \frac{P_{i,t}}{P_t}$:

$$\int_{0}^{1} p_{i,t}^{1-\sigma} = 1$$  \hspace{1cm} (2.3)

**Intermediate Goods Firms** There are a continuum of intermediate goods, each produced by a single monopolist. Each monopolist intermediate good producer receives revenue from selling goods to the final goods producer. The monopolist also faces idiosyncratic real marginal costs of production $MC_{i,t}$. We know that the demand of the final goods producer is given by equation (C.1).

An intermediate goods firm $i$ faces CES demand for its good each period and its profits are given by:

$$\Omega_{i,t} = \frac{P_{i,t}Y_{i,t}}{P_t} - MC_{i,t}Y_{i,t}$$  \hspace{1cm} (2.4)

s.t.

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\sigma} Y_t$$

We define idiosyncratic productivity so that we can express the idiosyncratic marginal cost of the firm as an aggregate marginal cost and idiosyncratic relative productivity i.e. $MC_{i,t} = \frac{MC_{i,t}}{a_{i,t}}$.

We input the CES demand into equation (2.4) to get equation (2.5).

$$\Omega_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{1-\sigma} Y_t - \frac{MC_t}{a_{i,t}} \left(\frac{P_{i,t}}{P_t}\right)^{-\sigma} Y_t$$  \hspace{1cm} (2.5)

We see that a firm’s real profits in a given period are a function $\Omega$ of a firm’s relative price and relative productivity $a_{i,t}$:

$$\Omega_{i,t} = \Omega(p_{i,t}, a_{i,t}) = Y_t \left[ p_{i,t}^{1-\sigma} - \frac{MC_t}{a_{i,t}} p_{i,t}^{-\sigma} \right]$$

\footnote{For example, if production is a constant function of labor, $Y_{i,t} = A_{i,t}L_t$ then the idiosyncratic marginal cost would be $MC_{i,t} = \frac{W_t}{P_t} \frac{1}{A_{i,t}}$ where $\frac{W_t}{P_t}$ is the real wage so we could rewrite the idiosyncratic marginal cost as $MC_{i,t} = \frac{MC_{i,t}}{a_{i,t}}$ where $MC_t = \frac{W_t}{P_t} \frac{1}{A_t}$, $a_{i,t} = \frac{A_{i,t}}{A_t}$.
**Profit Share**  We also note that since the final goods firm is competitive, total real profits \( \Omega_t \) are given by the sum of all intermediate firm real profits \( \Omega_{i,t} \):

\[
\Omega_t = \int_0^1 \Omega_{i,t} = \int_0^1 \frac{P_{i,t} Y_{i,t}}{P_t} - MC_t Y_{i,t}
\]

We can simplify this to yield:

\[
\Omega_t = Y_t - Y_t MC_t \nu_t \quad (2.6)
\]

where:

\[
\nu_t = \int_0^1 \left( \frac{P_{i,t}}{P_t} \right)^{-\sigma} \quad (2.7)
\]

From equation (2.6), we see that the profit share (the share of output \( Y_t \) given to owners of firms) is given by:

\[
\frac{\Omega_t}{Y_t} = 1 - MC_t \nu_t \quad (2.8)
\]

### 2.2.2 Calvo Pricing

We have the same basic final and intermediate good setup as in section 2.2.1. The key feature of Calvo pricing is that firms can only change their price with fixed probability \( \lambda \) each period. We denote the discount factor of the firm \( i \) periods into the future by \( M_{t,t+i} \) and the price the firm chooses when it does get to change its price as \( P_t^* \). We assume constant productivity for simplicity. Inputting the demand from the final goods firm (equation (C.1)) yields the following maximization problem:

\[
\max_{P_t^*} \mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j Y_{t+j} \left( \frac{P_{t+j}}{P_t} \right)^{\sigma-1} \left[ \left( \frac{P_t^*}{P_t} \right)^{1-\sigma} - MC_{t+j} \frac{P_{t+j}}{P_t} \left( \frac{P_t^*}{P_t} \right)^{-\sigma} \right] \right]
\]
Taking FOCs:

\[
\mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j Y_{t+j} \left( \frac{P_{t+k}}{P_t} \right)^{\sigma-1} \left[ (1 - \sigma) \left( \frac{P^*_t}{P_t} \right)^{-\sigma} + \sigma MC_{t+j} \frac{P_{t+j}}{P_t} \left( \frac{P^*_t}{P_t} \right)^{-\sigma-1} \right] \right] = 0
\]  

(2.9)

Simplifying equation (2.9):

\[
\mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j Y_{t+j} \left( \frac{P_{t+k}}{P_t} \right)^{\sigma-1} \left[ \frac{P^*_t}{P_t} - \frac{\sigma}{\sigma-1} MC_{t+j} \frac{P_{t+j}}{P_t} \right] \right] = 0
\]  

(2.10)

**Steady State**  
The steady state stochastic factor of the firm is defined as \( \beta \) i.e. \( \bar{M}_{t,t+i} = \beta^i \). Gross inflation is denoted as \( \Pi \) i.e. \( \Pi_{t+1} = \frac{P_{t+1}}{P_t} \). In steady state, the FOC (equation (2.10)) can be rewritten as:

\[
\sum_{k=0}^{\infty} (1 - \lambda)^k \beta^k \bar{\Pi}^{k\sigma} \left[ \left( \frac{P^*_t}{P_t} \right) \frac{1}{\bar{\Pi}^k} - \frac{\sigma}{\sigma-1} MC \right]
\]

This can be further simplified to get:

\[
\overline{MC} = \frac{\sigma - 1}{\sigma} \frac{1 - (1 - \lambda) \beta \bar{\Pi}^\sigma}{1 - (1 - \lambda) \beta \bar{\Pi}^{\sigma-1}} \left( \frac{P^*_t}{P} \right)
\]  

(2.11)

We can rewrite the price dispersion equation (equation (2.7)) recursively as:

\[
\nu_t = \lambda \left( \frac{P^*_t}{P_t} \right)^{-\sigma} + (1 - \lambda) \nu_{t-1} \Pi_t^\sigma
\]  

(2.12)

In steady state, equation (2.12) becomes:

\[
\bar{\nu} = \frac{\lambda}{1 - (1 - \lambda) \Pi^\sigma} \left( \frac{P^*_t}{P} \right)^{-\sigma}
\]
Multiplying both sides of equation (A.6) by $\bar{\nu}$ yields:

$$MC\bar{\nu} = \frac{\sigma - 1}{\sigma} \frac{1 - (1 - \lambda)\beta \Pi^{*\sigma}}{1 - (1 - \lambda)\beta \Pi^{*\sigma - 1}} \left( \frac{P^*}{P} \right)^{1-\sigma} \frac{\lambda}{1 - (1 - \lambda)\Pi^\sigma} \quad (2.13)$$

Next, we note that equation (C.2) can be rewritten recursively as:

$$1 = \lambda \left( \frac{P^*_t}{P_t} \right)^{1-\sigma} + (1 - \lambda)\Pi_{t}^{\sigma - 1} \quad (2.14)$$

In steady state equation (2.14) becomes:

$$\left( \frac{P^*_t}{P} \right)^{1-\sigma} = \frac{1}{\lambda} [1 - (1 - \lambda)\Pi^{\sigma - 1}] \quad (2.15)$$

Inputting equation (A.9) into equation (A.8) and simplifying yields:

$$MC\bar{\nu} = \frac{\sigma - 1}{\sigma} \left[ \frac{1 - (1 - \lambda)\Pi^{\sigma - 1}}{1 - (1 - \lambda)\beta \Pi^{\sigma - 1}} \right] \left[ \frac{1 - (1 - \lambda)\beta \Pi^{\sigma}}{1 - (1 - \lambda)\Pi^\sigma} \right] \quad (2.16)$$

This allows us to compute the profit share by equation (2.8).

### 2.2.3 Rotemberg Pricing

We have the same basic final and intermediate good setup as in section 2.2.1. The key feature of Rotemberg pricing is that firms have to pay a convex adjustment cost to update their price from whatever price they set in the previous period. Thus, unlike Calvo, they can change their price every period but have to pay a cost when they do so. The cost is the square of the log change in prices multiplied by the cost factor $\mu$ and output. Multiply the cost by output is a common assumption which helps to simplify the Math. We assume constant productivity for simplicity. Inputting the demand from the final goods firm (equation (C.1))
yields the following maximization problem:

$$\max_{\{P_t^*\}_{k=0}^\infty} \mathbb{E}_t \left[ \sum_{k=0}^\infty M_{t,t+k} Y_{t+k} \left( \left( \frac{P_{t+k}^*}{P^*_{t+k}} \right)^{1-\sigma} - MC_{t+k} \left( \frac{P_{t+k}^*}{P^*_{t+k}} \right)^{-\sigma} \right) - 0.5 \mu Y_t \log \left( \frac{P_t^*}{P^*_{t-1}} \right)^2 \right]$$

Taking FOCs yields:

$$[P_t^*]: Y_t P_t^\sigma (1-\sigma) P_t^\sigma - \sigma P_t^\sigma M C_t P_t^\sigma - \mu Y_t \frac{1}{P_t^*} \log \left( \frac{P_t^*}{P_t^*} \right) + \mu \mathbb{E}_t \left[ M_{t,t+1} Y_{t+1} \frac{1}{P_t} \log \left( \frac{P_{t+1}^*}{P_{t+1}^*} \right) \right] = 0$$

Multiply by $P_t$ and applying $P_t = P_t^*$:

$$\mu Y_t \log(\Pi_t) = Y_t (1-\sigma) + \sigma M C_t + \mu \mathbb{E}_t [M_{t,t+1} Y_{t+1} \log(\Pi_{t+1})] \quad (2.17)$$

We assume that $M_{t,t+1} = \beta \frac{Y_t}{Y_{t+1}}$. This is the correct stochastic discount factor to use when all output is consumed and agents have log utility since then $M_{t,t+1} = \beta^t \frac{C_{t+1}}{C_t}$. In this case, equation (2.17) simplifies to:

$$\mu \log(\Pi_t) = \sigma M C_t - (\sigma - 1) + \beta \mu \mathbb{E}_t [\log(\Pi_{t+1})] \quad (2.18)$$

Applying steady state to equation (2.18) yields:

$$\sigma M C - (\sigma - 1) = (1 - \beta) \mu \log(\bar{\Pi}) \quad (2.19)$$

### 2.2.4 Menu Costs

We have the same basic final and intermediate good setup as in section 2.2.1. They key feature of menu costs is that if firms update their price, they have to pay a cost $\mu$. If they do update their price, they pick a new optimal relative price $p_{i,t}^*$. If they do not change their price their absolute price is unchanged but, in relative terms, the price will fall (assuming positive inflation) by $\bar{\Pi}$ in relative terms.
We represent the problem agents face as a value function iteration in equation (2.20). The states are the relative price of the firm in the preceding period \( p_{i,t-1} \) and their current relative productivity \( a_{i,t} \). Each period, they choose one of two options: Firstly, they can reset their price and pay cost \( \mu \) (the first option in the curly bracket). Secondly, they can not change their price in which case their relative price changes by \( \bar{\Pi} \) (the second option in the curly bracket). This problem can be solved by standard value function iteration.

\[
V(p_{i,t-1}, a_{i,t}) = \max \left\{ \max_{p^*_i,t} \left[ \Omega(p^*_i,t, a_{i,t}) - \mu + \beta \mathbb{E}_t[V(p^*_i,t+1, a_{i,t+1})] \right], \Omega \left( \frac{p_{i,t-1}}{\Pi^*}, a_{i,t} \right) + \beta \mathbb{E}_t[V \left( \frac{p_{i,t-1}}{\Pi^*}, a_{i,t+1} \right)] \right\}
\]  

(2.20)

Partial equilibrium: If we take \( MC \) as given (the marginal cost is part of the profit function): Using the value functions computed, we can then compute the optimal reset price and the price ranges under which a firm will change their price for each productivity. This allows us to get the distribution of prices chosen by all firms.

General equilibrium: Under general equilibrium, we can no longer take \( MC \) as given and must solve for it. To do this, we keep solving for the distribution of prices under \( MC \) taking into account the level of inflation \( \bar{\Pi} \) until we find the \( MC \) that aggregates relative prices to 1 according to equation (2.3).

### 2.2.5 Parameterization

Each period represents one month. The baseline value for \( \beta \) is 0.96\(^{\frac{1}{12}}\) which is standard. The elasticity of substitution between varieties within industries (\( \sigma \)) is set to be \( \sigma = 8 \). This is in between the lower and upper bounds of respective 6 and 10 used in Carvalho et al. (2016).

For the standard Calvo model, we set \( \lambda = 0.1 \). This is in the range of the monthly frequency of price changes including price changes due to substitutions between different sectors found by Nakamura and Steinsson (2008).

The Rotemberg model yields identical log-linearised results as the Calvo model when inflation is zero - see appendix B.2.2. We therefore set \( \mu = 69.15 \) by equation (B.5) which
Table 2.2: Relationship between Inflation and the Profit Share under Alternative Price Rigidities Summary Table

<table>
<thead>
<tr>
<th>Inflation (p.p.)</th>
<th>0</th>
<th>2</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calvo</td>
<td>12.50</td>
<td>12.44</td>
<td>12.35</td>
</tr>
<tr>
<td>Rotemberg</td>
<td>12.50</td>
<td>12.44</td>
<td>12.36</td>
</tr>
<tr>
<td>Menu Cost Model</td>
<td>12.50</td>
<td>12.34</td>
<td>12.23</td>
</tr>
</tbody>
</table>

means that the Rotemberg pricing matches the Calvo pricing case around zero inflation.

For the menu cost model, we set $\mu = 0.0407$. This is calibrated so that the monthly frequency of price changes equals 8.7%. This matches the frequency of price changes, not including price changes due to substitutions, in Nakamura and Steinsson (2008). The reason we do not include price changes due to substitutions, unlike in the Calvo parameterization, is because the logic of a menu costs model is based upon an argument in Nakamura and Steinsson (2008) that price changes are only made because it is optimal for firms to reset their price. We set the idiosyncratic productivity to be an AR(1) process $\log(A_{i,t}) = \rho_a \log(A_{i,t-1}) + \epsilon_{i,t}, \epsilon_{i,t} \sim N(0, \sigma_a^2)$ where $\rho_a = 0.66, \sigma_a = 0.0428$ which are parameters taken from a partial equilibrium menu cost model in Nakamura and Steinsson (2008). This process is modeled using the method of Tauchen with 30 grid points.

2.3 Determinants of the Long Run Relationship between Inflation and Profit Share

2.3.1 Form of Price Rigidity

Figure 2.1 displays how a rise in inflation affects profitability under different price rigidities. The results are summarized in table 2.2. We observe that under standard price rigidities, a rise in inflation lowers profitability.

Theorem 1 demonstrates that under Calvo pricing when inflation rises, the profitability of
firms always falls as long as firms discount the future. A firm’s problem is to set its optimal markup (price over marginal cost). Without price rigidity, a firm’s optimal markup is just a tradeoff between maximizing demand and profits on each good sold. Under Calvo pricing, firms also have to tradeoff how the markup they set when they get to change their price will change if they do not get to reset their price in future periods. With higher inflation, firms should set a higher markup when they get to change their price since under higher inflation, a firm’s markup is likely to fall by more (since marginal costs rise more quickly) before firms can next reset their price. However, firms do not want to set a high markup in the current period since they care more about current profits than future profits due to discounting. Therefore, they do not raise their markups a lot in the current period meaning that on average firms’ markups fall.

**Theorem 1** (Calvo Pricing: Inflation and the Profit Share). *Under Calvo pricing with $\beta < 1$, raising inflation lowers the profit share.*
Theorem 2 demonstrates that under Rotemberg pricing when inflation rises, the profitability of firms always falls as long as firms discount the future. Under Rotemberg pricing, firms face convex adjustment costs of updating their price. If inflation rises, it requires firms to raise their price by more each period and therefore pay a higher marginal adjustment costs each period. Firms discount the future so they prefer to put off paying the adjustment cost until subsequent periods. Therefore, since a rise in inflation causes firms to delay raising their price by more, the average price firms set over the marginal cost (their markup) falls when inflation rises.

**Theorem 2** (Rotemberg Pricing: Inflation and the Profit Share). *Under Rotemberg pricing with $\beta < 1$, raising inflation lowers firm profitability.*

**Proof.** When $\beta < 1$, a rise in $\bar{\Pi}$ raises $\bar{MC}$ in appendix B.3.1. We know that there is no price dispersion since all firms act the same and thus $\bar{\nu} = 1$. Thus, we see that when $\bar{\Pi} \uparrow$, the profit share falls.

One difference we observe between Calvo and Rotemberg is that the impact of raising inflation on profitability gets stronger as inflation increases for Calvo but not for Rotemberg. This makes sense because, under Calvo, a rise in inflation worsens price dispersion by more under high inflation. Under Rotemberg, by contrast, there is no price dispersion and instead the marginal cost of the convex cost price adjustment increases approximately linearly ($\log(\bar{\Pi})$) in inflation as we see in equation (2.19).

With menu costs, a rise in inflation also means that firms need to raise their price by more each period as long as firms do not set their price more frequently. This would imply that, with discounting, a rise in inflation lowers average markups under similar logic to the Calvo case. However, since a firm’s pricing is state rather than time dependent, firms will choose to set their price more frequently. This will counteract the aforementioned Calvo effect. Consequently, we might expect that the effect of raising inflation on profitability will
be weaker under menu costs. We observe the opposite is true in figure 2.1.

One potential reason why inflation lowers the profit share by more is that a firm will not want to raise its price in line with inflation because there is a risk that they will get an idiosyncratic shock (for example a positive productivity shock) that would cause them to want to set a lower price in the future. In this case, they would have to pay two menu costs from raising and then lowering their price when they could have paid no menu cost by leaving their price unchanged. We would not get this effect in Calvo where firms pay no cost to update their price or in Rotemberg where firms prefer to make incremental changes to their price each period due to the convex nature of their cost of updating.

A second potential reason why inflation lowers the profit share by more than Calvo is that we parameterized the menu cost model so that the frequency of prices matches the data under 2% inflation. With inflation below 2% prices will update less frequently than this. Therefore, firms potentially set prices for longer and thus set lower markups on average (due to the standard Calvo reasoning) than in Calvo itself. I intend to further investigate the exact causes of the shape of the relationship between inflation and the profit share under menu costs.

2.3.2 Firm Discounting

In figure 2.2 the degree to which a fall in trend inflation from 2% to 0% lowers profitability given the level of firm discounting is plotted under different price rigidities. With a higher firm discount, firms care less about the future so they set a lower markup in response to a rise in inflation. Consequently, a rise in inflation lowers the profitability of firms by more. We can see this in figure 2.2.

We see that for Calvo and Rotemberg as $\beta$ approaches 1, a fall in inflation stops having an impact upon profitability. This makes sense because the reason why raising inflation lowers

---

2Additionally, the parameter for frequency of price changes in Calvo was based upon the empirically observed frequency of price change including price substitutions, unlike menu costs. Therefore, at 2% inflation, the frequency of price changes is slightly higher under Calvo anyway.
the markup and profitability relies upon firms discounting the future. We observe that a fall in inflation raises profitability by less under menu costs but that there is still a large effect when firms do not discount. This is because the additional reason for the relationship between inflation and profitability under menu costs (that firms worry about having to lower their price after raising them) does not require firm discounting to hold.

### 2.3.3 Elasticity of Substitution Between Goods

A higher elasticity of substitution between goods means that the demand for an intermediate firm’s good varies more when their price relative to their competitors changes. In figure 2.3, the degree to which a fall in inflation from 2% to 0% lowers profitability given different elasticities of subsitution between types of good is plotted for different price rigidities. We see that as $\sigma$ increases the effect gets stronger in Calvo but weaker in Rotemberg and menu costs.
In Rotemberg pricing, this implies that the marginal benefit of getting closer to the optimal markup rises so firms are happier to raise their price despite the high costs of price adjustment in the short-term. In menu cost pricing, a higher $\sigma$ implies that firms will change their price more frequently to get closer to the optimal markup. With more frequent price adjustment, the degree of monetary non-neutrality falls, reducing the impact of raising inflation on profitability. That being said, it should be noted that if firms set their price more frequently, the menu cost needed to generate the correct amount of price changes each period would be higher, which would boost the impact of raising inflation on profitability as is discussed in appendix B.3.3.

In Calvo, however, the frequency with which firms change their price is fixed so we do not get the same effect as in menu cost models. The costs of price dispersion increase when $\sigma$ rises because if there is a difference in firm prices, the final goods firm buys more from the cheaper firm, meaning that less is produced from a given amount of inputs across firms and
that profitability falls. There is much more price dispersion in the Calvo model than under
other price rigidities since whether firms get to change their price is not state-dependent.
Thus, when $\sigma$ is higher, a fall in inflation raises profitability by more in Calvo since a fall in
inflation lowers price dispersion by more.

### 2.4 Full Model

We also allow for an extension of section 2.2.2 with multiple sectors.

**Final Goods Firm** I allow for different industries with different weights and degrees of
price rigidity. This is done for two reasons. The primary reason is that allowing for different
degrees of price rigidities increases the degree of monetary non-neutrality which is otherwise
unrealistically low. See Carvalho (2006) for a detailed discussion. It is also more realistic to
allow for different industries with different degrees of price rigidity.

There is a single competitive final goods firm which aggregates goods in different indus-
tries to produce a final good. There are $J$ industries in total, denoted $1, \ldots, J$. The final
goods firm has CES production and each industry has a weight $a_j$ in production:

$$\left( \sum_{j=1}^{J} a_j^{\frac{1}{\sigma_j^2}} Y_j \frac{1}{\sigma_j^2} dj \right)^{\frac{\sigma_j^2}{\sigma_j^2 - 1}} = Y_t$$

Therefore, the final goods firm has the usual CES demand (taking into account industry
weights) for each industry good given by equation (2.21). The price aggregator also takes
the usual form given by equation (2.22). Note that weights $a_j$ need to be added for each
industry.

$$Y_{j,t} = a_j \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma_j} Y_t$$

$$P_t = \left( \sum_{j=1}^{J} a_j P_{j,t}^{1-\sigma_j} dj \right)^{\frac{1}{1-\sigma_j}}$$
Industry Aggregator  A perfectly competitive firm of firm $j$ aggregates all the intermediate goods in that industry to produce the good for sector $j$. The sector firm has the following production function:

$$Y_{j,t} = \left( \int_0^1 Y_{i,j,t}^{\frac{2}{\sigma}-1} di \right)^{\frac{\sigma}{\sigma-1}}$$

Therefore, the industry aggregator has the usual CES demand for each intermediate good given by equation (2.23). The price aggregator also takes the usual form given by equation (2.24).

$$Y_{i,j,t} = Y_{j,t} \left( \frac{P_{i,j,t}}{P_{j,t}} \right)^{-\sigma}$$ (2.23)

$$P_{j,t} = \left( \int_0^1 P_{i,j,t}^{1-\sigma} di \right)^{\frac{1}{1-\sigma}}$$ (2.24)

Intermediate Firm Profit Maximisation  Firms in each industry $j$ have a $\lambda_j$ probability of updating their price each period. When they do get to change their price, firms maximise equation (2.25) subject to the demand for their good from the industry aggregator firm (equation (2.23)). Firms discount future real profits by a fixed amount $\beta$. Intermediate firms therefore face the problem given in equation (2.25).

$$\max_{P_{j,t},Y_{i,j,t}} \mathbb{E}_t \left[ \sum_{k=0}^{\infty} M_{t,t+k} (1 - \lambda_j)^k \left[ \frac{P_{j,t}^{*} Y_{i,j,t+k}}{P_{t+k}} - MC_{j,t+k} Y_{i,j,t+k} \right] \right]$$ (2.25)

Like in the case with only one sector, we can aggregate the intermediate firms’ profits to get aggregate profits:

$$\Omega_t = Y_t - Y_t MC_t \nu_t$$ (2.26)

where we have industry price dispersion and total price dispersion terms:

$$\nu_t = \sum_{j=1}^{J} a_j \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma_2} \nu_{j,t} dj$$ (2.27)

$$\nu_{j,t} = \int_0^1 \left( \frac{P_{i,j,t}}{P_{j,t}} \right)^{-\sigma} di$$ (2.28)
Equation \((2.28)\) can be rewritten as equation \((2.29)\). Equation \((2.24)\) can be rewritten as equation \((2.30)\).

\[
\nu_{j,t} = \lambda_j \left( \frac{P^*_{j,t}}{P_{j,t}} \right)^{-\sigma} + (1 - \lambda_j)\nu_{j,t-1} \Pi_{j,t}^\sigma
\]

\[(2.29)\]

\[
1 = \lambda_j \left( \frac{P^*_{j,t}}{P_{j,t}} \right)^{1-\sigma} + (1 - \lambda_j)\Pi_{j,t}^{\sigma-1}
\]

\[(2.30)\]

\[
1 = \sum_{j=1}^J a_j \left( \frac{P_{j,t}}{P_t} \right)^{1-\sigma_2}
\]

\[(2.31)\]

We proceed by similar steps to the single sector Calvo case to derive the steady state. The steps are given in full in appendix B.4.1.

**Parameterization**

I set \(\beta = 0.96\) in the standard parameterization, as in section 2.2. The elasticity of substitution between firms within an industry is set to be 8 like in section 2.2. The elasticity of substitution between industries (\(\sigma_2\)) is set to 1 as in Shamloo (2010).\(^3\)

### 2.5 Results of Full Model

We now analyze the full model allowing for firm heterogeneity through multiple sectors which was outlined in section 2.4. The results under the baseline discount factor and the low factor are plotted against the results from the Calvo model with no heterogeneity in figure 2.4. Table 2.3 summarizes these results. We see that allowing for multiple sectors implies that raising inflation lowers profitability by more. A low firm discount factor, meaning that firms care less about the future, magnifies these results even further.

Average CPI inflation in the US in 1983-1992 and 2008-2017 was respectively 3.81% and 1.69%. The profit share (corporate profits before tax and without IVA and CCAAdj / GDP) in the US in 1983-1992 and 2008-2017 was respectively 5.86% and 9.26%. Therefore, we want to look at how much the fall in inflation from 3.81% to 1.69% can explain the rise

\(^3\)I actually set it to be 1.001 otherwise I would have to rewrite the indices since 1 is a special limiting case.
Figure 2.4: Long Run Relationship between Inflation and the Profit Share allowing for Firm Heterogeneity

![Graph showing the relationship between inflation and profit share with different scenarios.]

Table 2.3: Relationship between Inflation and the Profit Share under Firm Heterogeneity

<table>
<thead>
<tr>
<th>Summary Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation (p.p.)</td>
</tr>
<tr>
<td>Calvo</td>
</tr>
<tr>
<td>Calvo Multi-Sector</td>
</tr>
<tr>
<td>Calvo Multi-Sector with Low Firm Discount</td>
</tr>
</tbody>
</table>
in the profit share from 5.86% to 9.26%. We see in table 2.3 that once we allow for firm heterogeneity the firm profit share would have fallen by 0.47p.p. due to the fall in the profit share. In this case, the fall in inflation explains 14% of the rise in the profit share from 5.86% to 9.26%.

In practice, however, there is evidence that firms discount by more than the discount factor of the household. This makes sense since firms face financial frictions that imply that their cost of borrowing (their Weighted Average Cost of Capital (WACC)) is higher than the risk free rate. Jagannathan et al. (2016) estimated that it was 8% in 2003 when the expected ten year rate on real bonds ($r_e$) was 2.8p.p. Graham and Harvey (2011) estimated it was 10.0% in 2011Q1 when $r_e$ was 2.2%. Graham and Harvey (2012) then estimated it was 9.3% in 2012Q2 when $r_e = 1.3\%$. From these three surveys, the average wedge between WACC and the expected risk free real rate is 7p.p. which would imply that $\beta_f = \frac{0.96}{1.07} = 0.897$. I use this in an alternative parameterization. It is also worth noting that these surveys also measure actual firm discounting which is found to be even higher than their WACC would imply. If we use this lower level of firm discounting, we find that there would be a fall of 1.00p.p. in the profit share when trend inflation falls from 3.81% to 1.69%. In this case, the fall in inflation explains 29% of the rise in the profit share.

2.6 Empirical Evidence

In this section I present reduced form empirical evidence for the negative relationship between trend inflation and firm profits. There is existing empirical evidence for this. Bénabou (1992) finds evidence for this using reduced form regressions on US data. Banerjee and Russell (2001) find such a relationship using structural vector error correction models.

We can measure the markup indirectly through the labor share. Let’s assume that output is a linear function of labor (equation (2.32)). Then we know that the marginal cost firms
face will just equal the wage divided by productivity (equation (2.33)).

\[ Y_{i,t} = A_t L_{i,t} \]  \hspace{1cm} (2.32)

\[ MC_t = \frac{W_t}{A_t} \]  \hspace{1cm} (2.33)

Integrating over equation (C.1) yields equation (2.34). Inputting equation (2.34) into equation (2.32) yields equation (2.35). Combining equation (2.33) and equation (2.35) yields equation (2.36)

\[ \nu_t Y_t = \int_0^1 Y_{i,t} di \]  \hspace{1cm} (2.34)

\[ \nu_t Y_t = A_t L_t \]  \hspace{1cm} (2.35)

\[ MC_t \nu_t = \frac{W_t L_t}{Y_t} \]  \hspace{1cm} (2.36)

From equation (2.36), we observe that a 1p.p. rise in the labor share implies a 1p.p. fall in the firm profit share.

Equation (2.37) shows the regression relationship. It uses the same basic panel data structure as for the inflation-real rate relationship.

\[ \text{LaborShare}_{i,t} = \alpha_i + \delta_t + \beta \text{Inflation}_{i,t} + \Gamma \text{Controls}_t + u_{i,t} \]  \hspace{1cm} (2.37)

The panel is limited to OECD members. Annual data is used. Long-run inflation (\(\text{Inflation}_{i,t}\)) is measured as the moving average of the current and previous four years of CPI inflation.\(^4\) The real interest rate is measured by the 10 year expected return on safe bonds. The 10 year real rate is used since there is more data availability and its likely to be a much less noisy measure of the equilibrium real rate. To measure the 10 year real interest rate...

\(^4\)It may seem strange that the regressor is not the inflation target but nearly all inflation targets have not changed since they were introduced so the inflation target would be almost completely captured by the country fixed effects (\(\alpha_i\)).

\(^5\)Varying the measure to a different moving average does not appear to impact the results.
Table 2.4: Empirical Estimates of Relationship between Long Run Inflation and Long Run Markup

<table>
<thead>
<tr>
<th>LaborShare&lt;sub&gt;i,t&lt;/sub&gt;</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation&lt;sub&gt;i,t-4,t&lt;/sub&gt;</td>
<td>0.209*</td>
<td>0.304***</td>
<td>0.281***</td>
<td>0.298</td>
</tr>
<tr>
<td></td>
<td>(0.083)</td>
<td>(0.037)</td>
<td>(0.032)</td>
<td>(0.178)</td>
</tr>
<tr>
<td>country dummies</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>N</td>
<td>813</td>
<td>813</td>
<td>813</td>
<td>721</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

rate the measure of long-run inflation is subtracted from the nominal interest rate on 10 year government bonds. 6 I also allow for business cycle controls. 7 The business cycle controls are set to be GDP growth and change in unemployment at $t$ and $t - 1$.

The labor share is measured as the percentage net value added in production that is received as compensation by employees. I compute this for firms only since we are interested in measuring the markup which would be most related to firms’ labor share. This data is taken from National Accounts at the UN and the OECD. 8

The results are given in table 2.4. They do not vary very much once country fixed effects are introduction. A 1p.p. rise is associated with a 0.28p.p. rise in the markup. Once controls are added, the results are not significant due to a large increase in the standard error but I do not think this is concerning since the coefficient itself is unchanged and we are primarily interested in the third column regression without controls.

I verify these relationships are robust. Table B.1 verifies the relationships continue to

6Computing the measure of 10 year real interest rates by subtracting current inflation (rather than the measure of long-run inflation) from the nominal interest rate on 10 year government bonds yields similar results.

7It is undesirable to have controls that capture the long-run state of the economy since these could interfere in the long-run relationship between inflation and the real rate. Business cycle controls are short-term and should not generate this problem.

8Both the OECD and the UN National Accounts data for firms has some gaps. I take the UN data by default and fill in gaps with OECD data.
hold with just OECD members that joined before 1975 (these regressions exclude a number of mostly Eastern European countries that joined from the 1990s onwards). Table B.2 verifies the relationships still apply under low inflation. Table B.3 looks at whether the result remain for the period before 2000 only. Table B.4 analyzes whether the relationship continues to hold during/after 2000 only. Spurious regressions are generally considered to be less of a problem in panel data since we can control for common trends across countries by time fixed effects and idiosyncratic trends within countries are unlikely to drive results. However, I verify in table B.5 that the results still hold after differencing.

2.7 Conclusion

I have demonstrated that this relationship holds under non-Calvo forms of price rigidity. Indeed, the relationship can be stronger under menu costs, perhaps due to the incentive to hold off on raising prices in case there are future shocks that make a lower price more attractive or because a menu cost implies potentially lower price-setting frequency than Calvo under low inflation. I have also demonstrated that the size of the effect is stronger when firms discount the future more and can increase or decrease when the elasticity of substitution rises depending upon the price rigidity. I have shown that, once we incorporate firm heterogeneity, the fall in inflation from 3.82% to 1.69% would imply a rise of 0.47p.p. in the profit share which would account for 14% of the rise in firm profits in the US. If a more realistic degree of discounting by firms based upon the finance literature is allowed for then the fall in inflation would imply a rise of 1.00p.p. in the profit share which would account for 29% of the rise in firm profits in the US. I have also provided reduced form empirical evidence for the existence of such a relationship. In summary, the negative relationship between inflation and profits does appear to be general and of significance.
Chapter 3

A Long-Term Behavioral New Keynesian Model
3.1 Introduction

To resolve paradoxes and increase realism, economic models increasingly incorporate behavioral features. The dominant paradigm in macroeconomics, rational expectations, relies upon the strong assumption that agents fully understand the economy in which they live and can process information costlessly. Increasingly, economists are keen to relax the rational expectations assumption for realism and to try to explain current paradoxes.

A popular behavioral framework is sparse dynamic programming. Sparse dynamic programming was introduced in Gabaix (2014). It assumes that agents’ beliefs about shocks are partly rational but also partly behavioral. The behavioral part of agents’ expectations are based upon some default model. Typically, the default model is just the steady state of the model.

Assuming agents have a default model poses two problems. The first is that if there are persistent deviations away from the default model then agents’ expectations will be persistently wrong. For example, if agents base their consumption upon expectations of future income and that income rises persistently above its default then agents will persistently under-consume and never learn to raise their consumption enough. The second problem is that agents in sparse dynamic programming need to have a very good idea of what the default model is. Typically, therefore, they need to know the exact steady state of the economy in which they live. This is a strong assumption for a model that attempts to relax the amount of information that agents need to know.

My first main contribution is to introduce an alternative form of sparse dynamic programming based upon long-run updating which overcomes these problems. This framework assumes that the behavioral part of agents’ expectations is taken from their current beliefs which update every period based upon their new information. In this case, if there is a persistent change to a variable then agents’ behavioral expectations about that variable will slowly update to the new value. For example, if agents base consumption upon their ex-
expectations of future income and there is a persistent rise in income above the default then initially agents will under-consume but their behavioral beliefs will update meaning that eventually they consume the correct amount. This assumption also means that agents no longer need to know steady state. Instead, they just need to be able to have a rough idea of the average value that a variable tends to take based upon recent observations.

My second main contribution is to apply this alternative form of sparse dynamic programming to the New Keynesian model to derive a long-run Behavioral New Keynesian model. I assume that the representative agent no longer perfectly views output and the real interest rate which affects their consumption allocation problem. Otherwise, the New Keynesian features remain standard. Gabaix (2018) presented a Behavioral New Keynesian model based upon standard sparse dynamic programming. The key qualitative difference between the standard New Keynesian model, the standard Behavioral New Keynesian model and my model is the IS curve. In the non-behavioral IS curve, current output gaps and future output gaps have a 1:1 relationship. In the standard behavioral IS curve, future output gaps are discounted due to agents’ imperfect ability to perceive future consumption. In my behavioral IS curve, future output gaps are also discounted but there is an additional term that captures agents’ behavioral expectations which can update over time to re-establish to 1:1 relationship.

My third main contribution is to demonstrate that allowing updating to the behavioral expectations significantly affects the properties of the Behavioral New Keynesian model. In the standard Behavioral New Keynesian model, fixed interest rates can be determinate and stable and the costs of the zero lower bound can be small and bounded. This matches the recent experience of Japan where inflation appears to be pinned down and the Japanese economy does not seem to have suffered dramatically despite Japan being stuck at the zero lower bound for 25 years. However, once the behavioral expectations are allowed to update in sparse dynamic programming, fixed interest rates are never determinate and stable and the costs of the zero lower bound are increasing and unbounded. These properties are the
same as in the standard New Keynesian model without updating. Therefore, by allowing for updating to the behavioral beliefs, the properties of the behavioral model revert to more standard properties and the paradoxes are re-established.

The form of the IS curve drives whether a model is determinate under a fixed nominal interest rate. In the standard New Keynesian model, a rise in future output causes current output to rise due to consumption smoothing and due to the fact that a rise in future output raises expectations of future inflation from the New Keynesian Phillips Curve and thus lowers the real interest rate. Thus, if current output rises, future output also rises but by progressively less over time until we return back to steady state. We see that any current output is possible and thus we see we have indeterminacy. This effect can only be counteracted by raising the real interest rate through a more than one-to-one increase in the nominal interest rate in response to inflation. In the standard Behavioral New Keynesian model, however, we do not necessarily find these effects because part of agents’ expectations of future output is fixed to be steady state consumption. Therefore, a rise in current output requires a larger rise in future output and thus an explosive path. Consequently, we can get determinacy under a fixed interest rate rule. Once we allow for long-run updating in the Behavioral New Keynesian model, however, a rise in output means that expectations of output rise in the long-run. This implies that if current output rises, future output eventually returns to steady state and does not follow an explosive path so we have indeterminacy once again.

The form of the IS curve also determines the costs of remaining at the zero lower bound. In the standard Behavioral New Keynesian model, at the zero lower bound there is negative output but the degree to which output can fall is constrained by the fact that part of the agent’s expectations of future output is always the steady state. Therefore, output can fall by a limited amount. Once we allow for updating, however, the agent’s expectations fall when current output is below their expectations. Thus, over time, since the agent’s expectations of future output get worse, current output falls by more and more and the costs of the zero
lower bound are thus unbounded.

A related set of papers that this paper clearly relates to is Gabaix’s framework of sparse
dynamic programming. Gabaix (2014) introduced the basic idea of sparse dynamic pro-
gramming. Gabaix (2017) discussed how this could be applied broadly to Macroeconomics.
Gabaix (2018) applied this to the New Keynesian model to generate the Behavioral New
Keynesian model.

My paper also relates to a wider literature studying the impact of introducing behavioral
features into a New Keynesian model by other methods. One such paper is Woodford (2018).
Woodford considers a different behavioral framework where agents are rational up to some
finite period and then use a value function to capture their utility for the rest of time.
Woodford shows that applying this approach to the standard New Keynesian framework
with a fixed interest rate (or more generally a less than one-to-one response of the nominal
interest rate to inflation) yields instability in the long-run but not a multiplicity of solutions.
This is not identical to my results where I find indeterminacy and thus a multiplicity of
solutions. However, it is of a similar flavor in that both of the models imply that remaining
at the zero lower bound in a New Keynesian framework with behavioral features should
produce instability, unlike Gabaix (2018).

More broadly, this paper relates to recent papers on the discounted Euler equation. Typ-
ically, we expect that if consumers raise their consumption in the future by 1p.p. then
they will consumption smooth and raise their consumption by 1p.p. in the present ceteris
paribus. Euler discounting allows for the possibility that consumers don’t raise their con-
sumption today as much in response to future consumption. Gabaix (2018) introduces the
discount by effectively fixing a portion of the future expectations of consumption. McKay
et al. (2017) also introduce a discounted Euler equation model. The discounting there comes
from the fact that in the future with some probability agents will receive a fixed amount of
consumption. Broadly, a discounted Euler equation relies upon fixing some component of
future consumption or expectations of future consumption to be constant. This makes a lot

82
of sense in the short-run or when things change on a one-off basis like in forward guidance. However, in the long-term, this is not an innocuous assumption.

In section 2, I review the differences between the method Gabaix uses to solve for long-term expectations and the method I use in the context of a simple consumption allocation problem. In section 3, I define the alternative form of sparse dynamic programming which allows for long-run updating. I then introduce the New Keynesian model setup in section 4 before discussing the alternative Behavioral New Keynesian models that result from different behavioral assumptions in section 5. In sections 6 and 7, I look at respectively how allow for long-run updating affects determinacy and the costs of the zero lower bound. Section 8 concludes.

3.2 Simple Example

In this section, we consider how sparse dynamic programming affects a simple partial equilibrium consumption-allocation problem. We demonstrate that the implications do not appear reasonable in the long-term. We propose an alternative form of sparse dynamic programming which yields the same results as sparse dynamic programming in the short-term but does not suffer from the flaws of sparse dynamic programming in the long-term.

3.2.1 Setup

An agent lives for infinite periods. Each period they consume $C_t$ which yields utility $u(C_t)$. They maximise their lifetime utility and have a discount rate $\beta$. Agents receive income $Y_t$ each period. Any income that agents do not consume is saved for the next period and is denoted $S_{t+1}$. Agents start each period with savings they made in the last period $S_t$ on which they have received a fixed return of $r$. We allow for behavioral expectations, denoted $E^b$. Each period, they solve the following problem to compute their consumption today $C_t$.
and the amount they save for the next period $S_{t+1}$:

$$\max_{\{S_{t+i+1}, C_{t+i}\}_{i=0}^{\infty}} \mathbb{E}_t \left[ \sum_{i=0}^{\infty} \beta^i u(C_{t+i}) \right]$$

s.t.

$$S_{t+i+1} = Y_{t+i} - C_{t+i} + (1 + r) S_{t+i} \quad (3.1)$$

The first order condition of the problem is:

$$u'(C_{t+i}) = \beta(1 + r) \mathbb{E}_{t+i}[u'(C_{t+i+1})]$$

We assume that $r$ is such that $\beta(1 + r) = 1$. We make two approximations which mimic the effects of linearising without going through those steps. The first approximation we make is that $\mathbb{E}_{t+i}[u(C_{t+i+1})] \approx \frac{1}{\mathbb{E}_{t+i}[C_{t+i+1}]}$ which holds approximately under log utility. This yields:

$$\mathbb{E}_{t+i}[C_{t+i+1}] = C_{t+i} \quad (3.2)$$

Iterating over the budget constraint yields the lifetime budget constraint:

$$\sum_{i=0}^{\infty} \frac{C_{t+i}}{(1 + r)^i} = \sum_{i=0}^{\infty} \frac{Y_{t+i}}{(1 + r)^i} + (1 + r) S_t \quad (3.3)$$

We take expectations at $t$ of the lifetime budget constraint (equation (3.3)). On the left hand side, we input equation (3.2). On the right hand side, we define the log-linear deviation of income $\hat{Y}_t = \log(Y_t) - \log(\bar{Y})$ where $\bar{Y}$ is steady state income. We apply a second approximation that $\mathbb{E}_t[Y_{t+i}] \approx \bar{Y} \exp(\mathbb{E}_t[\hat{Y}_{t+i}])$. This yields:

$$\sum_{i=0}^{\infty} \frac{C_t}{(1 + r)^i} = \bar{Y} \sum_{i=0}^{\infty} \frac{\exp(\mathbb{E}_t[\hat{Y}_{t+i}])}{(1 + r)^i} + (1 + r) S_t$$
Simplifying this further yields:

\[ C_t = \frac{r}{1 + r} \bar{Y} \sum_{i=0}^{\infty} \frac{\exp(\mathbb{E}_t[\hat{Y}_{t+i}])}{(1 + r)^i} + rS_t \]  

(3.4)

In this paper, we are interested in what happens when we have long-run deviations from the steady state. We assume that:

\[ \hat{Y}_{t+i} = 0 \quad \forall t < 0 \]

\[ \hat{Y}_{t+i} = \log(2) \quad \forall t \geq 0 \]

In words, until period 0, we have been in steady state. From 0 onwards, income will be double its steady state value. Note that this implies \( \hat{Y}_{t+i} = 0 \forall t < 0 \) and \( \hat{Y}_{t+i} = \log(2) \forall t \geq 0 \).

### 3.2.2 Rational Expectations Solution

Applying \( \hat{Y}_{t+i} = \log(2) \) to equation (3.4) and simplifying yields:

\[ C_t = 2\bar{Y} + rS_t \]

Inputting this value for consumption into the individual period budget constraint (equation (3.1)) gives:

\[ S_{t+i} = S_t \]

We observe that consumption and savings are constant. We plot this in figure 3.1. The parameters are set as follows: \( r = 0.05, \bar{Y} = 1, S_0 = 0 \).

These results make intuitive sense. If an agent knows that they will receive higher income \( x \) every period they will consume \( x \) more every period so there will be no change in their savings over time. This is an application of the permanent income hypothesis (Friedman 1957).
3.2.3 Sparse Dynamic Programming Beliefs

The basic principle of sparse dynamic programming is that agents have imperfect expectations due to information processing costs. Since they do not want to expend energy trying to formulate exactly what will realisations of variables will be, they rely partly upon a default idea of what these variables will be. In [Gabaix (2014)], Gabaix demonstrates that we can sparse dynamic programming problems by two steps: 1. the agent works out how much attention to attribute to each variable; 2. the behavioral agent solves out the model given their imperfect attention. However, in later work [Gabaix (2017, 2018)], Gabaix assumes that agents have particular forms of inattention (for which their attention costs are not specified) which produces relatively simple behavioral macroeconomic models.

We take the latter approach. In particular, we assume that an agent only comprehends $M$ of the deviation in a variable $i$ periods in the future from steady state (where $0 \leq M \leq 1$).

In this case:

$$E_t[\hat{Y}_{t+i}] = M E_t[\hat{Y}_{t+i}] \forall i \geq 0$$

Since we know that $\hat{Y}_{t+i} = \log(2) \forall i$ then:

$$E_t[\hat{Y}_{t+i}] = M \log(2) \forall i \geq 0$$
Figure 3.2: Consumption Allocation Problem under Standard Sparse Dynamic Programming

(a) Consumption

(b) Savings

Note that here we have assumed that even in the current period, agents do not perfectly observe $\hat{Y}_t$ which helps to keep the math simple but is not important.

3.2.4 Sparse Dynamic Programming Solution

Applying behavioral expectations from our simple version of sparse dynamic programming implies that the solution for consumption (equation (3.4)) becomes:

$$C_t = \frac{r}{1 + r} \hat{Y} \sum_{i=0}^{\infty} \frac{\exp(M \log(2))}{(1 + r)^i} + rS_t$$

Simplifying yields the following processes for consumption and savings:

$$C_t = 2^M \hat{Y} + rS_t$$

$$S_{t+1} = [2 - 2^M] \hat{Y} + S_t$$

$M = 1$ is a special case where we agents perfectly observe future variables and we move back to the rational expectations case. Otherwise, when $0 \leq M < 1$, we observe that consumption is always less than income and savings rises indefinitely. We plot the path of consumption and savings in figure 3.2. $M$ is set to be 0.5.
The intuition here is that the agent again tries to apply the permanent income hypothesis and consume the same every period. However, they underestimate income in exactly the same way forever. They anticipate that they will have an income each period of $2^M\bar{Y}$ when actually each period going forward their income is $2\bar{Y}$.

To see why we get this result, we can rewrite the agent’s behavioral expectations as follows:

$$E^b_t[\hat{Y}_{t+i}] = ME_t[\hat{Y}_{t+i}] + (1-M)\hat{Y}$$  \hspace{1cm} (3.5)

The agent places $M$ weight upon the correct expectations $\hat{Y}_{t+i}$ and $1-M$ weight upon the steady state of $\hat{Y}_{t+i}$ (denoted with a bar). Since $\hat{Y}_{t+i}$ is the log deviation from steady state of $Y_{t+i}$ its steady state value is just 0. Therefore, when an agent has imperfect behavioral expectations of a variable and there are persistent deviations in a particular direction away from steady state, the behavioral expectations will be biased towards the steady state. In our case, income is persistently higher than steady state so the agent underestimates income and therefore consumes too little and saves too much indefinitely.

### 3.2.5 Long-Term Sparse Dynamic Programming Beliefs

In practice, there are two problems with the behavioral expectations assumptions we have made. Firstly, we have assumed that even if there are persistent deviations away from steady state, agents never learn. Secondly, we assume that agents perfectly perfectly know what the steady state of the model is. To avoid these issues, we introduce an alternative form of behavioral expectations:

$$E^b_t[\hat{Y}_{t+i}] = ME_t[\hat{Y}_{t+i}] + (1-M)\mu_{\hat{Y},t}$$

where:

$$\mu_{\hat{Y},t} = \alpha_{\hat{Y}}\hat{Y}_{t-1} + (1-\alpha_{\hat{Y}})\mu_{\hat{Y},t-1}$$
We assume that agents base the behavioral part of their expectations of $\hat{Y}_t$ upon some long-run estimate of $\hat{Y}$ which they update slowly. Therefore, if there is a permanent increase in $\hat{Y}$, the behavioral expectations will be biased downwards initially but will slowly rise to take into account the permanent increase.

Moreover, this framework does not require agents to know steady state exactly. Instead, they only need to estimate $\hat{Y}$ based upon recent observations, which appears more reasonable. Of course, if the economy has been in steady state for a long time, the expectations of $\hat{Y}$ will naturally converge to the steady state value of $\hat{Y}$.

3.2.6 Long-Term Sparse Dynamic Programming Solution

Applying behavioral expectations from our simple version of sparse dynamic programming implies that the solution for consumption (equation (3.4)) becomes:

$$C_t = \bar{Y} \exp(M \log(2) + (1 - M)\mu_{\hat{Y},t}) + rS_t$$

(3.6)

Therefore, their savings process is given by:

$$S_{t+1} - S_t = \bar{Y}[2 - \exp(M \log(2) + (1 - M)\mu_{\hat{Y},t})]$$

(3.7)

Since we have been in steady state until $0$ ($\hat{Y}_t = 0 \forall t < 0$), the agent’s long-term expectations of $\hat{Y}_t$ prior to $0$ must be $0$ ($\mu_{\hat{Y},0} = 0$). The path of consumption and savings under long-term sparse dynamic programming is given in figure 3.3 $\alpha$ is set to be $0.1$.

Immediately after the permanent rise in income, the agent’s long-term expectations of $\hat{Y}_t$ are approximately zero so the agent believes that approximately $\hat{Y}_t = M \log(2)$, as in the sparse dynamic programming case. Consequently, initially, the agent saves a positive amount, as in normal sparse dynamic programming, because they underestimate the value of future income.

However, over time, the long-term expectations of $\hat{Y}$ update to take into account the
persistent deviation in $\hat{Y}$. In the long-term their expectations therefore converge to the updated value of $\hat{Y} = \log(2)$. Formally, $\lim_{t\to\infty} \mu_{\hat{Y},t} = \log(2)$. This implies that agents eventually come to believe correctly that their income has doubled and thus, in the long-term, they consume this amount more and their savings stop increasing.

### 3.3 Long-Term Sparse Dynamic Programming

In this section, we formally define the alternative form of sparse dynamic programming I introduce where the irrational part of agents’ expectations slowly updates. I also give the standard sparse dynamic programming case for comparison.

We give the definition in the case of a standard constrained maximization setup. However, it would be similarly defined in other cases. Agents maximise some utility function $u$ each period. They do this for $T$ periods which is potentially infinite and discount each period by $\beta$. Agents maximize over some control $a_{t+i}$ give the process $F$ by which a state $x_{t+i}$ updates. $u$ and $F$ are both potentially affected by shocks $\epsilon_{t+i}$ each period. Therefore, a rational agent
would characterize the problem as follows:

\[
\max_{\{a_{t+i}\}_{t=0}^T} \mathbb{E}_t \left[ \sum_{i=0}^T \beta^i u(a_{t+i}, x_{t+i}, \epsilon_{t+i}) \right]
\]

s.t.

\[
x_{t+i+1} = F(a_{t+i}, x_{t+i}, \epsilon_{t+i})
\]

Note that we allow for multiple controls, states and shocks so \(a_{t+i}, x_{t+i}, \epsilon_{t+i}\) are all vectors.

**Definition 3** (Sparse Dynamic Programming Approach). In standard sparse dynamic programming, agents maximize the following problem each period to solve for \(a_t\) and \(x_{t+1}\):

\[
\max_{\{a_{t+i}\}_{t=0}^T} \mathbb{E}_t \left[ \sum_{i=0}^T \beta^i u(a_{t+i}, x_{t+i}, m_i \circ \epsilon_{t+i}) \right]
\]

s.t.

\[
x_{t+i+1} = F(a_{t+i}, x_{t+i}, m_i \circ \epsilon_{t+i})
\]

\(m_i\) is a vector of the degree to which agents understand deviations in \(\epsilon_{t+i}\) from steady state. \(\circ\) represents the element-by-element multiplication i.e. \((a, b) \circ (c, d) = (ac, bd)\).

Gabaix (2014) describes in more detail the justification and derivation of \(m_i\). However, in later work, Gabaix takes the degree to which agents understand deviations as given. We follow the latter approach here.

**Definition 4** (Long-Run Sparse Dynamic Programming Approach). In long-run sparse dynamic programming, agents maximize the following problem each period to solve for \(a_t\) and \(x_{t+1}\):

\[
\max_{\{a_{t+i}\}_{t=0}^T} \mathbb{E}_t \left[ \sum_{i=0}^T \beta^i u(a_{t+i}, x_{t+i}, m_i \circ \epsilon_{t+i} + (1 - m_i) \mu_{t+i}) \right]
\]

s.t.

\[
x_{t+i+1} = F(a_{t+i}, x_{t+i}, m_i \circ \epsilon_{t+i} + (1 - m_i) \mu_{t+i})
\]
where:

\[ \mu_{\epsilon,t} = \alpha_{\epsilon} \epsilon_{t-1} + (1 - \alpha_{\epsilon}) \mu_{\epsilon,t-1} \]

### 3.4 Simple New Keynesian Model Setup

Here we describe the setup for the simple New Keynesian model that we will consider, under different behavioral expectations, for the rest of the paper.

#### 3.4.1 Households

There is a representative agent with behavioral expectations, denoted with a \( b \) superscript. They maximize their lifetime utility where \( u(C) \) is their utility from consumption and they discount by \( \beta \). Agents receive income \( Y_t \) each period. There is a nominal bond market which pays off gross nominal interest of \( I_{t-1} \) each period (denoted with a \( t-1 \) since this rate was chosen by the central bank in the previous period). Therefore, agents face the following consumption-allocation problem:

\[
\max \mathbb{E}^b_t \left[ \sum_{i=0}^{\infty} \beta^i u(C_{t+i}) \right]
\]

s.t.

\[
C_t + \frac{B_{t+1}}{P_t} = Y_t + \frac{B_t}{P_t} I_{t-1} \tag{3.8}
\]

We derive a standard Euler equation (except with Behavioral expectations):

\[
u'(C_t) = \beta \mathbb{E}_t [u'(C_{t+1}) R_{t+1}] \tag{3.9}\]

where:

\[
R_{t+1} = \frac{I_t}{\Pi_{t+1}} \tag{3.10}
\]

\[
\Pi_{t+1} = \frac{P_{t+1}}{P_t}
\]
The agent receives income from working and from the profits of the firm. Agents can work for real wage $W_t$. The amount the agent works is denoted $L_t$. They get disutility $v(L_t) = \frac{1}{\eta}L_t^{1+\eta}$ from working. Therefore, working for $\epsilon$ units provides utility of $W_t u'(C_t)\epsilon$ at a cost $v'(L_t)$. Thus, by a perturbation argument, we have the consumption-labor condition:

$$W_t u'(C_t) = v'(L_t)$$  \hspace{1cm} (3.11)

Log-linearising equation (3.11) yields:

$$\hat{W}_t - \frac{1}{\gamma}\hat{C}_t = \eta\hat{L}_t$$  \hspace{1cm} (3.12)

### 3.4.2 Firms

We use the standard New Keynesian firm setup with monopolistic intermediate goods firms producing with linear labor-only production functions and facing Calvo price rigidities, and a CES competitive firm aggregating the intermediate goods. By standard computations, which are relegated to appendix [C.1.1], the New Keynesian Phillips Curve is given by:

$$\pi_t = \kappa x_t + \beta E_t[\pi_{t+1}]$$  \hspace{1cm} (3.13)

where $\pi_t$ is log-linearised gross inflation ($\pi_t = \hat{\Pi}_t$) and $x_t$ represents the output gap:

$$x_t = \hat{Y}_t - \hat{Y}^n_t$$  \hspace{1cm} (3.14)

We could have introduced behavioral agents on the firm side like in Gabaix’s behavioral New Keynesian model but all this really does is lower the $\beta$ parameter. It doesn’t qualitatively change the resulting model.
3.4.3 Monetary and Fiscal Policy

We assume the central bank pursues a simple Taylor Rule where $i_t$ represents log-linearised gross nominal interest ($i_t = \hat{I}_t$):

$$i_t = \phi_\pi \pi_t$$  \hspace{1cm} (3.15)

In equilibrium, we set that there are no bonds so $B_t = 0$.

3.4.4 Equilibrium

We know that in equilibrium all output is consumed so:

$$Y_t = C_t$$  \hspace{1cm} (3.16)

Log-linearising equation (3.16) yields:

$$\hat{Y}_t = \hat{C}_t$$  \hspace{1cm} (3.17)

Note that even if there are no future shocks the path that agents decide at time $t$ may not be what they do at $t + 1$ because we allow for the possibility of non-rational expectations.

3.4.5 Rewriting the Household Problem

We assume CRRA utility i.e. $u(C_t) = C_t^{1-\frac{1}{\gamma}}$. We use the general notation that $\hat{X}_t = \log(X_t) - \log(\bar{X})$. We can then log-linearise equations (3.9) and (3.10) to yield:

$$E_t^h[\hat{C}_{t+1}] = \hat{C}_t + \gamma E_t^h[\hat{R}_{t+1}]$$  \hspace{1cm} (3.18)

$$\hat{R}_{t+1} = \hat{I}_t - \hat{\Pi}_{t+1}$$  \hspace{1cm} (3.19)

We define real bonds to be $b_t = \frac{B_t}{\hat{R}_{t-1}}$. We can rewrite the budget constraint (equa-
tion (3.8)) in real terms as:

$$C_t + b_{t+1} = Y_t + R_t b_t$$ (3.20)

Iterating over the real individual period budget constraint (equation (3.20)) yields the lifetime budget constraint:

$$\sum_{i=0}^{\infty} \frac{C_{t+i}}{\Pi_{j=1}^{i} R_{t+j}} = \sum_{i=0}^{\infty} \frac{Y_{t+i}}{\Pi_{j=1}^{i} R_{t+j}} + R_t b_t$$ (3.21)

Since there are no bonds in equilibrium, $b_t = 0$. We then take expectations at $t$ and log-linearise the lifetime budget constraint (equation (3.21)) to get:

$$\sum_{i=0}^{\infty} \frac{\bar{C}}{\bar{R}^i} (\hat{C}_{t+i} - \sum_{j=1}^{i} \hat{R}_{t+j}) = \sum_{i=0}^{\infty} \frac{\bar{Y}}{\bar{R}^i} (\hat{Y}_{t+i} - \sum_{j=1}^{i} \hat{R}_{t+j})$$ (3.22)

By equation (3.16), $\bar{C} = \bar{Y}$. We then take behavioral expectations at $t$ and substitute out $E_t[\hat{C}_{t+i}]$ with $\hat{C}_t$ using the Euler equation (equation (3.18)):

$$\sum_{i=0}^{\infty} \frac{1}{\bar{R}^i} (\hat{C}_t + (\sigma - 1) \sum_{j=1}^{i} E_t^b[\hat{R}_{t+j}]) = \sum_{i=0}^{\infty} \frac{1}{\bar{R}^i} (E_t^b[\hat{Y}_{t+i}] - \sum_{j=1}^{i} E_t^b[\hat{R}_{t+j}])$$ (3.23)

We can simplify to get\(^1\)

$$\hat{C}_t = \frac{\bar{R} - 1}{\bar{R}} \sum_{i=0}^{\infty} \frac{1}{\bar{R}^i} E_t^b[\hat{Y}_{t+i}] - \frac{\sigma}{\bar{R}} \sum_{i=0}^{\infty} \frac{E_t^b[\hat{R}_{t+i+1}]}{\bar{R}^i}$$ (3.24)

### 3.5 Behavioral IS Curve

In section 3.4, we derived the New Keynesian Phillips Curve (equation (3.13)) and defined a monetary policy curve (equation (3.15)). To get a full New Keynesian model, we need to derive an IS curve. We consider the derivation of the IS curve under different behavioral

\(^1\)To get this, note that $\sum_{i=0}^{\infty} \frac{1}{\bar{R}^i} \hat{C}_t = \frac{\bar{R}}{\bar{R} - 1} \hat{C}_t$ and $\sum_{i=0}^{\infty} \frac{1}{\bar{R}^i} \sum_{j=1}^{i} \hat{R}_{t+j} = \frac{\bar{R} - 1}{\bar{R}} \sum_{i=0}^{\infty} \frac{\hat{R}_{t+i+1}}{\bar{R}^i}$.
assumptions.

3.5.1 No Behavioral Features

In the rational case, $E_t[\hat{Y}_{t+i}] = E_t[\hat{R}_{t+i}]$. Then equation (3.24) becomes:

$$\hat{C}_t = \frac{R - 1}{R} \sum_{i=0}^{\infty} \frac{1}{R^i} E_t[\hat{Y}_{t+i}] - \frac{\sigma}{R} \sum_{i=0}^{\infty} E_t[\hat{R}_{t+i+1}]$$

(3.25)

We can rewrite this iteratively as:

$$\hat{C}_t = \frac{R - 1}{R} \hat{Y}_t - \frac{\sigma}{R} E_t[\hat{R}_{t+1}] + \frac{1}{R} E_t[\hat{C}_{t+1}]$$

(3.26)

By applying equation (3.17), we can simplify equation (3.26) to become:

$$\hat{C}_t = E_t[\hat{C}_{t+1}] - \sigma E_t[\hat{R}_{t+1}]$$

(3.27)

We define $r^n_t$ to be the ex-ante log-linearised natural real interest rate i.e. $r^n_t = E_t[\hat{R}^n_{t+1}]$. Applying equation (3.19) to equation (3.27) and then subtracting the natural version (the version without price rigidities) of equation (3.27) yields:

$$x_t = E_t[x_{t+1}] - \sigma(i_t - E_t[\pi_{t+1}] - r^n_t)$$

(3.28)

Equation (3.28) is just the standard IS equation. Without behavioral features, the New Keynesian model is then characterized by equations (3.13), (3.15) and (3.28).

3.5.2 Standard Sparse Dynamic Programming

In this section, we apply the same form of sparse dynamic programming as in Gabaix (2018). Effectively, this assumes that if a variable is observed $i$ periods from now, the agent only
captures $M^i$ of the deviation from steady state of this variable. Mathematically, we set:

$$E_t^b[\hat{Y}_{t+i}] = M^i E_t[\hat{Y}_{t+i}]$$

$$E_t^b[\hat{R}_{t+i+1}] = M^i E_t[\hat{R}_{t+i+1}]$$

Note that, even though agents have imperfect expectations, their expectations still partly depend upon rational expectations. One explanation is that the behavioral agents fully understand the structure of the world in which they live and construct rational forecasts taking into account their behavioral biases. However, it seems counterintuitive that agents can construct rational forecasts but possess behavioral expectations. A more reasonable explanation is that there exists a forecaster who understands the structure of the world, including the behavioral part of agents’ expectations, and is able to construct rational forecasts. Therefore, we are assuming that the behavioral agents partly take into account what the forecaster predicts but also partly base their expectations upon a default model.

Then equation (3.24) becomes:

$$\hat{C}_t = \bar{R} - \frac{1}{R} \sum_{i=0}^{\infty} M^i \bar{R} \hat{Y}_{t+i} - \sigma \bar{R} \sum_{i=0}^{\infty} M^i \bar{R} \hat{R}_{t+i+1}$$

(3.30)

We can rewrite this iteratively as:

$$\hat{C}_t = \frac{\bar{R} - 1}{\bar{R}} \hat{Y}_t - \frac{\sigma}{\bar{R}} \bar{R} \hat{R}_{t+1} + \frac{M}{\bar{R}} \bar{R} \hat{C}_{t+1}$$

(3.31)

By applying equation (3.17), we can simplify equation (3.31) to become:

$$\hat{C}_t = M E_t[\hat{C}_{t+1}] - \sigma E_t[\hat{R}_{t+1}]$$

(3.32)

Applying equation (3.19) to equation (3.32) and then subtracting the natural version of
equation $\text{(3.32)}$ yields the following IS curve:

$$x_t = M\mathbb{E}_t[x_{t+1}] - \sigma(i_t - \mathbb{E}_t[\pi_{t+1}] - r^n_t) \quad (3.33)$$

Equation $\text{(3.33)}$ is the same as equation $\text{(3.28)}$ except that the future output gap is discounted by $M$. With standard sparse dynamic programming, the New Keynesian model is then characterized by equations $\text{(3.13)}$, $\text{(3.15)}$ and $\text{(3.33)}$.

### 3.5.3 Long-Run Sparse Dynamic Programming

In this section, we apply the form of long-run sparse dynamic programming introduced in this paper. In particular, we set that the agent’s expectations of income are as follows:

$$\mathbb{E}_t^b[Y_{t+i}] = M^i\mathbb{E}_t[Y_{t+i}] + (1 - M^i)\mu_{Y,t}$$

where:

$$\mu_{Y,t} = \alpha_{Y}\hat{Y}_{t-1} + (1 - \alpha_{Y})\mu_{Y,t-1} \quad (3.34)$$

We only apply long-run sparse dynamic programming to income. We continue to assume that expectations of the real interest rate satisfy equation $\text{(3.29)}$. The reasons we do this is because it leads to easier Math and because this is all we need to get very different results to the standard Behavioral New Keynesian model in Gabaix (2018).

As in section $\text{3.5.2}$ agents partly base their expectations upon the true rational expectations. A reasonable explanation for where these rational expectations come from is that there exists a forecaster who understands the structure of the world, including the nature of agents’ expectations, and construct rational forecasts from this. Agents then partly base their expectations on these rational forecasts but also upon their long run beliefs.
Then equation (3.24) becomes:

$$\hat{C}_t = \frac{\bar{R} - 1}{R} \sum_{i=0}^{\infty} \frac{1}{R^i} (M^i \mathbb{E}_t[\hat{Y}_{t+i}] + (1 - M^i) \mu_{Y,t}) - \frac{\sigma}{R} \sum_{i=0}^{\infty} \frac{M^i}{R^i} \mathbb{E}_t[\hat{R}_{t+i+1}]$$  \hspace{1cm} (3.35)

Aggregating the $\mu_{Y,t}$ terms:

$$\hat{C}_t = \frac{\bar{R} - 1}{R} \sum_{i=0}^{\infty} \frac{M^i}{R^i} \mathbb{E}_t[\hat{Y}_{t+i}] + \frac{1}{R - 1} \mu_{Y,t} - \frac{\sigma}{R} \sum_{i=0}^{\infty} \frac{M^i}{R^i} \mathbb{E}_t[\hat{R}_{t+i+1}]$$  \hspace{1cm} (3.36)

Rewriting this iteratively:

$$\hat{C}_t = M \mathbb{E}_t[\hat{C}_{t+1}] + \frac{1 - M}{R - M} \mu_{Y,t} - M \frac{1 - M}{R - M} \mu_{Y,t+1} + \frac{\bar{R} - 1}{R} \hat{Y}_t - \frac{\sigma}{R} \mathbb{E}_t[\hat{R}_{t+1}]$$  \hspace{1cm} (3.37)

By applying equation (3.17), we can simplify equation (3.37) to become:

$$\hat{C}_t = M \mathbb{E}_t[\hat{C}_{t+1}] + \frac{1 - M}{R - M} \mu_{Y,t} - M \frac{1 - M}{R - M} \mathbb{E}_t[\mu_{Y,t+1}] - \sigma \mathbb{E}_t[\hat{R}_{t+1}]$$

Adding and subtracting $M \frac{1 - M}{R - M} \mu_{Y,t}$ from section 3.5.3 yields:

$$\hat{C}_t = M \mathbb{E}_t[\hat{C}_{t+1}] + (1 - M) \mu_{Y,t} - M \frac{1 - M}{R - M} \mathbb{E}_t[\mu_{Y,t+1}] - \mu_{Y,t} - \sigma \mathbb{E}_t[\hat{R}_{t+1}]$$  \hspace{1cm} (3.38)

Applying equations (3.14), (3.19) and (3.34) to equation (3.38) yields and then subtracting the natural version of equation (3.32) yields the following IS curve:

$$x_t = M \mathbb{E}_t[x_{t+1}] + (1 - M) \mu_{x,t} - M \frac{1 - M}{R - M} (\mathbb{E}_t[\mu_{x,t+1}] - \mu_{x,t}) - \sigma (i_t - \mathbb{E}_t[\pi_{t+1}] - r^n_t)$$  \hspace{1cm} (3.39)

where:

$$\mu_{x,t} = \alpha_Y x_{t-1} + (1 - \alpha_Y) \mu_{x,t-1}$$  \hspace{1cm} (3.40)

\[\text{To get this, note that } \frac{\bar{R} - 1}{R} \frac{1 - M^i}{R^i} = \frac{\bar{R} - 1}{R} \left( \frac{\bar{R}}{R - 1} - \frac{\bar{R}}{R - M} \right) = 1 - \frac{\bar{R} - 1}{R - M} = \frac{1 - M}{R - M}.\]
We can simplify this further by inputting equation (3.40) into equation (3.39):

\[
x_t = \tilde{M}E_t[x_{t+1}] + (1 - \tilde{M})\mu_{x,t} - \tilde{\sigma}(i_t - E_t[\pi_{t+1}] - \nu^n_t)
\]

(3.41)

where:

\[
\tilde{M} = \frac{M}{1 + \alpha \gamma M \frac{1-M}{R-M}}
\]

\[
\tilde{\sigma} = \frac{\sigma}{1 + \alpha \gamma M \frac{1-M}{R-M}}
\]

Equation (3.41) is similar to equation (3.33) except that it accounts for the slow updating of the irrational part of the agent’s expectations about future output. With long-run sparse dynamic programming, the New Keynesian model is then characterized by equations (3.13), (3.15), (3.40) and (3.41).

### 3.6 Determinacy in the Behavioral New Keynesian Model

#### 3.6.1 Standard Behavioral New Keynesian Model

Gabaix (2018) demonstrates it is possible to get determinacy when \( \phi_\pi < 1 \), including for a fixed interest rule, once we allow for standard sparse dynamic programming behavioral features. I briefly review these results.

**Theorem 5.** In the Behavioral New Keynesian model based upon standard sparse dynamic programming, \( \phi_\pi < 1 \) can yield determinacy and stability.

**Proof.** See appendix C.2.1.

Determinacy implies that there is a unique non-explosive expected path that can be followed by variables in a model. In other words, there is indeterminacy if we can pick any value for the current period output gap \( x_t \) without the model exploding. Stability implies that if we let the system run without shocks then it will converge to a non-explosive...
equilibrium.

In the rational case, when current output $x_t$ rises, expectations of the future output gap $E_t[x_{t+1}]$ do not rise by as much since higher future output means higher future inflation which (with a fixed interest rate) lowers the real interest rate and raises output today relative to the future. Therefore, we do not have determinacy (a unique equilibrium path) because when output can rise today away from steady state by any amount without future output then following an explosive path.

Once we allow for sparse dynamic programming i.e. $M < 1$, when $x_t$ rises, expectations of the future output gap can rise by more than the rise in the output gap today. This is because the only way agents will increase their output today is if it is rational to expect a larger increase in output in the future. This is an explosive path since it requires expectations of output to get larger further into the future. Thus, we see that allowing for behavioral features can generate determinacy.

We can also see this Mathematically by simplifying the model by setting $\beta = 0, \phi_x = 0, r_n = 0$ to obtain a simpler Phillips Curve and a fixed interest rate rule. Inputting these into equation (3.33) yields:

$$E_t[x_{t+1}] = \frac{1}{M + \sigma\kappa} x_t \quad (3.42)$$

We observe that when $M + \sigma\kappa < 1$ in equation (3.42), if $x_t \neq 0$ then $E_t[x_{t+1}]$ must explode. The only way current output is greater than 0 is if future output grows explosively. Therefore, in a non-explosive path, current output must be pinned down to 0.

It may seem strange that a rise of 1p.p. in output today can require a rise of more than 1p.p. in the future in the Behavioral New Keynesian model. However, Behavioral New Keynesian agents still conduct consumption smoothing. So, for agents to consume 1p.p. more today they still need to believe that their consumption will rise in the future. Since agents have limited attention to the future, they underestimate the degree to which output will rise in the future. Therefore, there needs to be a larger increase in future output than in the standard New Keynesian model to generate a 1p.p. increase in output today. If the
degree of inattention is large enough then it can be that a 1p.p. rise in output requires a more than 1p.p. rise in output in the future.

3.6.2 Long-Run Behavioral New Keynesian Model

Theorem 6 demonstrates that when the nominal interest rates responds less than one-to-one to inflation, the model must never be determinate and stable. This implies that an unreponsive nominal interest rate will not pin down inflation and thus that the economy will be unstable at the zero lower bound.

**Theorem 6.** In the Behavioral New Keynesian model based upon long-run sparse dynamic programming, $\phi_\pi \leq 1$ is never determinate and stable.

**Proof.** See appendix C.2.2.

To see why, we set $\beta = 0, \phi_\pi = 0, r^p_n = 0$ to obtain a simpler Phillips Curve and a fixed interest rate rule. Inputting these into equation (3.41) yields:

$$x_t = (\tilde{M} + \tilde{\sigma} \kappa)\mathbb{E}_t[x_{t+1}] + (1 - \tilde{M})\mu_{x,t}$$

where:

$$\mu_{x,t} = \alpha_{\tilde{Y}} x_{t-1} + (1 - \alpha_{\tilde{Y}})\mu_{x,t-1}$$

Therefore, to generate an increase in current output, future output would need to increase. However, over time, the agent’s behavioral expectations of output would rise. Therefore, to generate a rise in current output, output in the long-run is likely to need to rise by less than the rise in current output. Thus, a rise in current output is not likely to require the economy to follow an explosive path so the solution is indeterminate. This is similar to the standard New Keynesian model but very different to the behavioral New Keynesian model where a rise in current output requires the economy to follow an explosive path.
3.7 Costs of Zero Lower Bound in the Behavioral New Keynesian Model

We investigate the costs of being at the zero lower bound for a long period of time.

3.7.1 Standard Behavioral New Keynesian Model

We consider an experiment where the natural real rate of interest stays indefinitely at some constant level low enough to force the economy to remain at the zero lower bound. When \( r^n < -\pi^* \), it is necessary to set \( i < 0 \) to achieve stable inflation. This is not thought to be possible due to the zero lower bound (at least not significantly below zero). So we consider an experiment where \( r^n < -\pi^* \) indefinitely.

We discuss how to compute this experiment in appendix C.3.1. We can show that the costs of hitting the ZLB persistently can be bounded under standard sparse dynamic programming. A 1p.p. fall in \( \mathbb{E}_t[\hat{x}_{t+1}] \) ceteris paribus implies a \( M \)p.p. fall in \( \hat{x}_t \) where \( M < 1 \). Therefore, even if the real interest rate is positive, agents do not react fully to the future expected fall in the output gap. Therefore, the output gap can converge to some bounded level given in appendix C.3.1.

This is very different to the baseline non-behavioral New Keynesian model where, under rational expectations, the costs must intuitively be unbounded. A 1p.p. fall in \( \mathbb{E}_t[\hat{x}_{t+1}] \) ceteris paribus implies a 1p.p. fall in \( \hat{x}_t \). Therefore, if the real interest rises due to the central bank’s inability to lower the nominal interest rate and thus the real interest rate (which would boost current output), the output gap at \( t \) must be persistently non-negligibly lower than at \( t + 1 \). We know this will continue indefinitely and therefore the costs of remaining at the zero lower bound are unbounded.
3.7.2 Long-Run Behavioral New Keynesian Model

In the long-term Behavioral New Keynesian model, the costs of the zero lower bound are unbounded. We demonstrate this in appendix [C.3.2]

The output gap cannot converge to some stable level that bounds the costs of the zero lower bound, unlike in the standard Behavioral New Keynesian model. In the short-term, the output gap may converge to some relatively stable level since agents don’t respond fully now to output gaps in the future due to imperfect observation of the future. However, we cannot converge to this output gap completely because in the long-term the expectations of agents about their future income falls which lowers agents’ current output in a vicious cycle.

3.8 Conclusion

I introduce an alternative form of sparse dynamic programming that allows us to study how behavioral features affect models in the long-term. I do this by letting go the assumption in the Behavioral New Keynesian model that part of expectations always depend upon the steady state. I instead consider a framework in which the imperfect part of expectations within the behavioral framework slowly update according to agents’ observations. I apply this to the New Keynesian model. Within this setup, some of the key results of the Behavioral New Keynesian model reverse. It is no longer possible for a less than one-to-one response of nominal interest rates to inflation (including a fixed nominal interest rate rule) to generate determinacy and stability. This is the same as the New Keynesian model but the opposite of the Behavioral New Keynesian model. And the zero lower bound produces unbounded increasing costs. Again, this is the same as the New Keynesian model but the opposite of the Behavioral New Keynesian model. Therefore, the question of how to resolve the apparent paradoxes in the New Keynesian model at the zero lower bound remains open.
Bibliography


Appendix A

Appendix to Chapter 1
A.1 Intuition in a Simplified Model Details

In this section, additional details are provided relating to section 1.2.

A.1.1 Calvo Pricing: Inflation and the Markup

Here, the standard Calvo pricing setup is assumed and the relationship between inflation and the markup is derived. This is an example of the inflation-markup relationship that is discussed in section 1.2.1

There are a continuum of intermediate firms indexed by $i$ which produce differentiated intermediate goods. There is a competitive final goods firm with constant elasticity of substitution (CES) production with CES parameter $\sigma$. The final goods firm’s demand for good $i$ is represented by:

$$Y_{i,t} = \left(\frac{P_{i,t}}{P_t}\right)^{-\sigma} Y_t$$  \hspace{1cm} (A.1)

where:

$$\int_0^1 P_{i,t} Y_{i,t} di = P_t Y_t$$  \hspace{1cm} (A.2)

$$P_t = \left(\int_0^1 P_{i,t}^{1-\sigma} di\right)^{\frac{1}{1-\sigma}}$$  \hspace{1cm} (A.3)

$$Y_t = \left(\int_0^1 Y_{i,t}^{\sigma-1} di\right)^{\frac{\sigma}{\sigma-1}}$$

Intermediate firms have some probability $\lambda$ of updating their price each period. Firms have the following maximisation problem:

$$\max_{P_t^*,Y_{t+k}^*} \mathbb{E}[\sum_{k=0}^{\infty} (1 - \lambda)^k \beta^k \left[P_t^* Y_{t+k}^* - MC_{t+k} P_{t+k} Y_{t+k}^*\right]]$$  \hspace{1cm} (A.4)

s.t. $\forall k$:

$$Y_{t+k}^* = Y_{t+k} \left(\frac{P_t^*}{P_{t+k}}\right)^{-\sigma}$$  \hspace{1cm} (A.5)
Inputting equation (A.5) into equation (A.4) and taking FOCs yields:

\[\mathbb{E}_t \left[ \sum_{k=0}^{\infty} (1 - \lambda)^k \beta^k ((1 - \sigma) P_t^{\sigma-\sigma} Y_{t+k} P_t^{\sigma-1} + \sigma MC_{t+k} Y_{t+k} P_t^{\sigma-\sigma-1} P_{t+k}^\sigma) \right] \]

In steady state, this can be rewritten as:

\[\sum_{k=0}^{\infty} (1 - \lambda)^k \beta^k \bar{\Pi}^k \sigma \left( \frac{P^*_t}{P_t} \right)^{1 - \sigma} \left( \frac{P^*_t}{P_t} \right)^{1 - (1 - \lambda) \beta \bar{\Pi}^\sigma} \left( \frac{P^*_t}{P_t} \right)^{1 - (1 - \lambda) \beta \bar{\Pi}^{\sigma-1}} \left( \frac{P^*_t}{P_t} \right) \]

This can be further simplified to get:

\[MC = \frac{\sigma - 1}{\sigma} \frac{1 - \sigma}{1 - (1 - \lambda) \beta \bar{\Pi}^\sigma} \left( \frac{P^*_t}{P_t} \right) \]

(A.6)

A price dispersion term is defined:

\[\nu_t = \int_0^1 \left( \frac{P_{i,t}}{P_t} \right)^{1 - \sigma} di \]

(A.7)

This can be rewritten as:

\[\nu_t = \lambda \left( \frac{P^*_t}{P_t} \right)^{1 - \sigma} + (1 - \lambda) \nu_{t-1} \bar{\Pi}_t^\sigma \]

In steady state:

\[\bar{\nu} = \frac{\lambda}{1 - (1 - \lambda) \bar{\Pi}^\sigma} \left( \frac{P^*_t}{P_t} \right)^{1 - \sigma} \]

Multiplying both sides of equation (A.6) by \(\bar{\nu}\) yields:

\[\frac{MC \bar{\nu}}{\bar{\nu}} = \frac{\sigma - 1}{\sigma} \frac{1 - \sigma}{1 - (1 - \lambda) \beta \bar{\Pi}^\sigma} \left( \frac{P^*_t}{P_t} \right)^{1 - \sigma} \frac{\lambda}{1 - (1 - \lambda) \bar{\Pi}^\sigma} \]

(A.8)

Next, we note that equation (A.3) can be rewritten as

\[1 = \lambda \left( \frac{P^*_t}{P_t} \right)^{1 - \sigma} + (1 - \lambda) \bar{\Pi}_t^{\sigma-1} \]
where $\Pi_t = \frac{P_t}{P_{t-1}}$ and also $\pi_t = \Pi_t - 1$. In steady state:

$$
\left( \frac{P^*}{P} \right)^{1-\sigma} = \frac{1}{\lambda} [1 - (1 - \lambda)\bar{\Pi}^{\sigma-1}] \tag{A.9}
$$

Inputting equation (A.9) into equation (A.8) and simplifying yields:

$$
\bar{MC}\bar{\nu} = \frac{\sigma - 1}{\sigma} \left[ \frac{1 - (1 - \lambda)\bar{\Pi}^{\sigma-1}}{1 - (1 - \lambda)\beta\bar{\Pi}^{\sigma-1}} \right] \left[ 1 - (1 - \lambda)\beta\bar{\Pi}^{\sigma} \right] \tag{A.10}
$$

We set $\bar{m} = (\bar{MC}\bar{\nu})^{-1}$ which is a measure of the effective markup (this is explained in appendix A.1.2). Therefore, we can rewrite equation (A.10) as:

$$
\bar{m} = \frac{\sigma}{\sigma - 1} \left[ \frac{\bar{\Pi}^{1-\sigma} - (1 - \lambda)\beta}{\bar{\Pi}^{1-\sigma} - (1 - \lambda)} \right] \left[ \frac{\bar{\Pi}^{-\sigma} - (1 - \lambda)}{\bar{\Pi}^{-\sigma} - (1 - \lambda)\beta} \right] \tag{A.11}
$$

Raising $\bar{\Pi}$ increases the first square bracket but lowers the second. When $\beta < 1$, the second square bracket dominates since $\bar{\Pi}^{-\sigma}$ changes by more than $\bar{\Pi}^{1-\sigma}$ so raising average inflation lowers the markup.

By average markup, I’m effectively referring to the average markup weighted by sales. This is why I get qualitatively different results to Ascari and Sbordone (2014) who consider the non-weighted average markup and show that it can rise when inflation rises.

**Formal: Raising Inflation Lowers the Markup** The result that raising inflation lowers the markup when $\beta < 1$ is now shown formally. First derivatives are applied:

$$
d\bar{m} \quad \frac{d\bar{m}}{d\bar{\Pi}} = \left[ \frac{(1 - \sigma)\bar{\Pi}^{-\sigma}}{\bar{\Pi}^{1-\sigma} - (1 - \lambda)\beta} - \frac{(1 - \sigma)\bar{\Pi}^{-\sigma}}{\bar{\Pi}^{1-\sigma} - (1 - \lambda)} - \frac{\sigma\bar{\Pi}^{-\sigma-1}}{\bar{\Pi}^{-\sigma} - (1 - \lambda)} + \frac{\sigma\bar{\Pi}^{-\sigma-1}}{\bar{\Pi}^{-\sigma} - (1 - \lambda)\beta} \right] \left[ \frac{\bar{\Pi}^{1-\sigma} - (1 - \lambda)\beta}{\bar{\Pi}^{1-\sigma} - (1 - \lambda)} \right] \left[ \frac{\bar{\Pi}^{-\sigma} - (1 - \lambda)}{\bar{\Pi}^{-\sigma} - (1 - \lambda)\beta} \right]
$$

112
Concentrating on the first square bracket:

\[
\frac{d\tilde{m}}{d\bar{\Pi}} \propto -\frac{(1 - \sigma)\bar{\Pi}^{-\sigma}(1 - \lambda)(1 - \beta)}{(\bar{\Pi}^{1-\sigma} - (1 - \lambda)\beta)(\bar{\Pi}^{1-\sigma} - (1 - \lambda))} - \frac{\sigma\bar{\Pi}^{-\sigma-1}(1 - \lambda)(1 - \beta)}{(\bar{\Pi}^{-\sigma} - (1 - \lambda))(\bar{\Pi}^{-\sigma} - (1 - \lambda)\beta)}
\]

Rearranging and remove \((1 - \lambda)(1 - \beta)\) (which would not be possible if \(\beta = 1\)):

\[
\frac{d\tilde{m}}{d\bar{\Pi}} \propto \frac{\bar{\Pi}^{-\sigma}}{(\bar{\Pi}^{1-\sigma} - (1 - \lambda)\beta)(\bar{\Pi}^{1-\sigma} - (1 - \lambda))} - \frac{\sigma}{\sigma - 1} \frac{\bar{\Pi}^{-\sigma-1}}{(\bar{\Pi}^{-\sigma} - (1 - \lambda))(\bar{\Pi}^{-\sigma} - (1 - \lambda)\beta)}
\]

Substitute out \(\frac{\sigma}{\sigma - 1}\) using equation (A.11) and rearrange to get:

\[
\frac{d\tilde{m}}{d\bar{\Pi}} \propto \bar{\Pi}^{-\sigma}(\bar{\Pi}^{-\sigma} - (1 - \lambda))^2 - \tilde{m}\bar{\Pi}^{-\sigma-1}(\bar{\Pi}^{1-\sigma} - (1 - \lambda))^2
\]

Simplifying:

\[
\frac{d\tilde{m}}{d\bar{\Pi}} \propto \bar{\Pi}^{-\sigma}(\bar{\Pi}^{-\sigma} - (1 - \lambda))^2 - \tilde{m}\bar{\Pi}^{-\sigma-1}(\bar{\Pi}^{1-\sigma} - (1 - \lambda))^2
\]

\(\tilde{m} \geq 1\) otherwise firms exit the market and \(\frac{d\tilde{m}}{d\bar{\Pi}}\) is decreasing in \(\tilde{m}\). Therefore, \(\frac{d\tilde{m}}{d\bar{\Pi}}\) takes its highest possible value when \(\tilde{m} = 1\). I show that even in this case \(\frac{d\tilde{m}}{d\bar{\Pi}} < 0\) and thus the markup always decreases in inflation. Under \(\tilde{m} = 1\), simplify to yield:

\[
\frac{d\tilde{m}}{d\bar{\Pi}}_{\tilde{m}=1} \propto \bar{\Pi}^{-3\sigma} + (1 - \lambda)^2\bar{\Pi}^{-\sigma} - [\bar{\Pi}^{1-3\sigma} + \bar{\Pi}^{-\sigma-1}(1 - \lambda)^2]
\]

\[
\frac{d\tilde{m}}{d\bar{\Pi}}_{\tilde{m}=1} \propto (\bar{\Pi} - 1)((1 - \lambda)^2 - \bar{\Pi}^{-2\sigma})
\]

Then note that by equation (A.11) \(\bar{\Pi}^{-\sigma} \geq 1 - \lambda\) since this is needed to guarantee \(\tilde{m}\) is positive. Thus, \(\frac{d\tilde{m}}{d\bar{\Pi}} < 0\) so I have demonstrated that when average inflation rises, the average markup always falls under Calvo pricing when \(\beta < 1\).
A.1.2 Asset Demand with Price Dispersion

This section provides the derivation for the profit of the firm section 1.2.2 when prices are dispersed.

Intermediate firms produce output \( Y_{i,t} \) by a linear production function over labor \( L_{i,t} \):

\[
Y_{i,t} = L_{i,t}
\]  \hspace{1cm} (A.12)

Inputting equation (A.1) into equation (A.12) and aggregating yields:

\[
Y_t \nu_t = L_t
\]  \hspace{1cm} (A.13)

where \( \nu_t \) is defined by equation (A.7) and:

\[
L_t = \int_0^1 L_{i,t}
\]

The real marginal cost \( MC_t \) of intermediate firms is just their wage bill \( W_t \):

\[
MC_t = W_t
\]  \hspace{1cm} (A.14)

Each intermediate firm has the following real profits \( \Omega_{i,t} \):

\[
\Omega_{i,t} = \frac{P_t Y_{i,t}}{P_t} - MC_t Y_{i,t}
\]  \hspace{1cm} (A.15)

Inputting equation (A.1) into equation (A.15) and aggregating (where we also apply equation (A.2)) yields:

\[
\Omega_t = (1 - MC_t \nu_t) Y_t
\]  \hspace{1cm} (A.16)
where:
\[ \Omega_t = \int_0^1 \Omega_{i,t} di \]

Next, we define \( m_t \) to be \( \frac{1}{MC_u} \). We input this into equation (A.16):

\[ \Omega_t = (1 - \frac{1}{m_t}) Y_t \]  (A.17)

We observe that \( m_t \) represents the inverse of real costs of firms, which is a measure of the average markup across firms. When \( m_t = 1 \) (the competitive case), firms make no profits. However, as the markup rises \( m_t \uparrow \), firms make higher profits.

Next, we multiply the RHS of equation (A.17) by \( \frac{W_t}{MC_t} \) (which equals 1 by equation (A.14) and apply equation (A.13) before simplying:

\[ \Omega_t = (1 - \frac{1}{m_t}) \frac{W_t}{MC_t} \Omega_t \]

\[ \Omega_t = (m_t - 1) \frac{W_t}{P_t} L_t \]

We see that we get the same as the case without price dispersion.

### A.1.3 OLG in a Non-Annualized Model

Figure [A.1] shows the case where the OLG model is non-annualized, unlike figure 1.3 where it is annualized. The real interest rate is very high since it represents the return from one generation to the next.
A.2 Model Details

In this section, additional details are provided on the derivations of section 1.3.

A.2.1 Firms

Cost Minimisation Details

Intermediate firms minimise their costs. They face the following problem:

$$\min_{K_{i,j,t}, L_{i,j,t}} (1 + \tau)(r_t + \tau)K_{i,j,t} + W_tL_{i,j,t}$$

s.t.

$$Y_{i,j,t} = A_t K_{i,j,t}^\alpha L_{i,j,t}^{1-\alpha}$$
Note that $r_t, W_t$ are real variables. Setting up a Lagrangean and solving yields:

\[(1 + \tau)(r_t + \delta) = \lambda_{i,j,t} \alpha A_t K_{i,j,t}^{\alpha - 1} L_{i,j,t}^{1 - \alpha} \]  
(A.18)

\[(1 + \tau)W_t = \lambda_{i,j,t} (1 - \alpha) A_t K_{i,j,t}^\alpha L_{i,j,t}^{-\alpha} \]  
(A.19)

We can combine equations (A.18) and (A.19) to get:

\[\frac{K_{i,j,t}}{L_{i,j,t}} = \frac{\alpha}{1 - \alpha} \frac{W_t}{r_t + \delta} \]  
(A.20)

We observe in equation (A.20) that the ratio of each firm’s capital to labor \(\frac{K_{i,j,t}}{L_{i,j,t}}\) is determined by aggregate variables and is therefore the same across firms. Thus: \(\frac{K_{i,j,t}}{L_{i,j,t}} = \frac{K_t}{L_t}\). \(\lambda_{i,j,t}\) is the marginal cost after tax of production of the firm. This is constant for all firms and is equal to the real marginal cost after tax of the firm. \(MC_t\) is defined to be the marginal cost of firms before tax so that \(\lambda_{i,j,t} = (1 + \tau)MC_t\). This yields:

\[MC_t = \frac{r_t + \delta}{\alpha A_t K_t^{\alpha - 1} L_t^{1 - \alpha}} \]  
(A.21)

\[MC_t = \frac{W_t}{(1 - \alpha) A_t K_t^\alpha L_t^{-\alpha}} \]  
(A.22)

**Aggregation of Cost Minimisation Conditions Details**

Rewrite output as follows:

\[Y_{i,j,t} = A_t K_{i,j,t}^{\alpha} L_{i,j,t}^{1 - \alpha} \]

\[Y_{j,t} \left(\frac{P_{i,j,t}}{P_{j,t}}\right)^{-\sigma} = A_t \left(\frac{K_{i,j,t}}{L_{i,j,t}}\right)^\alpha L_{i,j,t} \]

Taking integrals and noting that the ratio \(\frac{K_{i,j,t}}{L_{i,j,t}}\) is the same across \(i, j\):

\[Y_{j,t} \nu_{j,t} = A_t \left(\frac{K_t}{L_t}\right)^\alpha L_{j,t} \]
\[ a_j \sigma_j \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma} \nu_{j,t} = a_j A_t \left( \frac{K_t}{L_t} \right)^\alpha L_{j,t} \]

Taking integrals again to get a condition with aggregate output:

\[ Y_t \int_0^1 a_j \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma} \nu_{j,t} dj = A_t \left( \frac{K_t}{L_t} \right)^\alpha L_t \]

Applying the definition of marginal costs and inputing the lump sum transfer, equation (1.49) can be rewritten more simply as equation (A.23). Also note that the lump sum transfer equals \( \tau (r_t K_{i,j,t} + W_t L_{i,j,t}) \).

\[ \Omega_{i,j,t} = \frac{P_{i,j,t} Y_{i,j,t}}{P_t} - (r_t + \delta) K_{i,j,t} + W_t L_{i,j,t} \]

\[ \Omega_{i,j,t} = \frac{P_{i,j,t} Y_{i,j,t}}{P_t} - MC_t Y_{i,j,t} \quad (A.23) \]

The same steps can then be followed with real profits:

\[ \Omega_{i,j,t} = \frac{P_{i,j,t} Y_{i,j,t}}{P_t} - MC_t Y_{i,j,t} \]

\[ \Omega_{j,t} = \frac{P_{j,t} Y_{j,t}}{P_t} - MC_t Y_{j,t} \nu_{j,t} \]

\[ \Omega_{j,t} = \frac{P_{j,t} Y_{j,t}}{P_t} - MC_t Y_{j,t} \nu_{j,t} \]

\[ \Omega_t = Y_t - MC_t a_j Y_t \left( \frac{P_{j,t}}{P_t} \right)^{-\sigma} \nu_{j,t} \]

By definition of \( \nu_t \):

\[ Y_t \nu_t = A_t K_t^\alpha L_t^{1-\alpha} \]

\[ \Omega_t = Y_t - Y_t MC_t \nu_t \]
Rewriting Price Evolution Equations Details

The evolution of the price dispersion equation for industry \( j \) (equation (2.29)) can be rewritten as:

\[
\nu_{j,t} = \int_{0}^{1} \left( \frac{P_{i,j,t}}{P_{j,t}} \right)^{-\sigma} \, di \\
= \lambda_{j} \left( \frac{P_{j,t}^{*}}{P_{j,t}} \right)^{-\sigma} + (1 - \lambda_{j}) \int_{0}^{1} \left( \frac{P_{i,j,t-1}}{P_{j,t}} \right)^{-\sigma} \, di \\
= \lambda_{j} \left( \frac{P_{j,t}^{*}}{P_{j,t}} \right)^{-\sigma} + (1 - \lambda_{j})\nu_{j,t-1}\Pi_{j,t}^{\sigma}
\]

The price aggregator for individual industries (equation (2.24)) can be rewritten as:

\[
P_{j,t}^{1-\sigma} = \int_{0}^{1} P_{i,j,t}^{1-\sigma} \, di \\
= \lambda_{j}P_{j,t}^{*1-\sigma} + (1 - \lambda_{j})P_{j,t-1}^{1-\sigma} \\
1 = \lambda_{j} \left( \frac{P_{j,t}^{*}}{P_{j,t}} \right)^{1-\sigma} + (1 - \lambda_{j})\Pi_{j,t}^{\sigma-1}
\]

Also observe that by definition:

\[
\Pi_{j,t} = \frac{P_{j,t}}{P_{j,t-1}} = \frac{P_{j,t}}{P_{t}} \frac{P_{t}}{P_{t-1}} \frac{P_{t-1}}{P_{j,t-1}}
\]

And the final price evolution equation (equation (2.22)) can be rewritten as:

\[
P_{t}^{1-\sigma_2} = \sum_{j=1}^{J} a_{j} P_{j,t}^{1-\sigma_2} \\
1 = \sum_{j=1}^{J} a_{j} \left( \frac{P_{j,t}}{P_{t}} \right)^{1-\sigma_2}
\]
Firm Price Maximisation Details

The firm’s problem is:

$$\max_{P^*_j,t, Y_{i,j,t}} \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \left( \frac{\beta_f}{R} \right)^k (1 - \lambda_j)^k \left[ \frac{P^*_j Y_{i,j,t+k}}{P_{t+k}} - (1 + \tau) MC_{j,t+k} Y_{i,j,t+k} \right] \right]$$ \quad (A.24)

s.t.

$$Y_{i,j,t+k} = \left( \frac{P^*_t}{P_{j,t+k}} \right)^{-\sigma} Y_{j,t+k}$$ \quad (A.25)

Equation (B.7) can be input into equation (B.8) to get:

$$\max_{P^*_j,t} \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \left( \frac{\beta_f}{R} \right)^k (1 - \lambda_j)^k \left[ P^*_j \left( 1 - \sigma \right) P_{j,t+k}^{-\sigma} Y_{j,t+k} \left( 1 + \tau \right) MC_{j,t+k} P^*_j P_{j,t+k}^{-\sigma} \frac{1}{P_{t+k}} Y_{j,t+k} \right] \right]$$ \quad (A.26)

Taking FOCs:

$$\mathbb{E}_t \left[ \sum_{k=0}^{\infty} \left( \frac{\beta_f}{R} \right)^k (1 - \lambda_j)^k \left[ (1 - \sigma) P_{j,t}^{-\sigma} P_{j,t+k} Y_{j,t+k} \left( 1 + \tau \right) MC_{j,t+k} P_{j,t}^{-\sigma} Y_{j,t+k} \right] \right]$$ \quad (A.27)

Rearranging:

$$\mathbb{E}_t \left[ \sum_{k=0}^{\infty} \left( \frac{\beta_f}{R} \right)^k (1 - \lambda_j)^k P_{j,t+k} Y_{j,t+k} \left( 1 + \tau \right) MC_{j,t+k} \right]$$ \quad (A.28)

Inputting $Y_{j,t+k}$ and dividing by $P_{j,t}$:

$$\mathbb{E}_t \left[ \sum_{k=0}^{\infty} \left( \frac{\beta_f}{R} \right)^k (1 - \lambda_j)^k \left( \frac{P_{j,t+k}}{P_{j,t}} \right)^{1-\sigma} \left( \frac{P_{j,t+k}}{P_{t+k}} \right)^{-\sigma} Y_{t+k} \left[ \frac{P_{j,t} P_{j,t+k} P_t}{P_{t+k}} - \frac{\sigma}{\sigma - 1} \left( 1 + \tau \right) MC_{j,t+k} \right] \right]$$ \quad (A.29)

Next, define:

$$U_{j,t} = \mathbb{E}_t \left[ \sum_{k=0}^{\infty} \left( \frac{\beta_f}{R} \right)^k (1 - \lambda_j)^k \left( \frac{P_{j,t+k}}{P_{j,t}} \right)^{1-\sigma} \left( \frac{P_{j,t+k}}{P_{t+k}} \right)^{-\sigma} Y_{t+k} \frac{P_t}{P_{t+k}} \right]$$
\[ V_{j,t} = \mathbb{E}_t \sum_{k=0}^{\infty} \left( \frac{\beta f}{R} \right)^k (1 - \lambda_j)^k \left( \frac{P_{j,t+k}}{P_{j,t}} \right)^\sigma \left( \frac{P_{j,t+k}}{P_{t+k}} \right)^{-\sigma} Y_{t+k} \frac{\sigma}{\sigma - 1} (1 + \tau) MC_{j,t+k} \]

Then:

\[ U_{j,t} \frac{P_{j,t}^*}{P_{j,t}} P_{j,t} - V_{j,t} = 0 \]

\[ U_{j,t}, V_{j,t} \] can then be rewritten:

\[ U_{j,t} = \left( \frac{P_{j,t}}{P_{t}} \right)^{-\sigma_2} Y_t + \mathbb{E}_t \left[ \frac{\beta f}{R} (1 - \lambda_j) \Pi^\sigma_{j,t+1} \Pi_{t+1}^{-1} U_{j,t+1} \right] \]

\[ V_{j,t} = \left( \frac{P_{j,t}}{P_{t}} \right)^{-\sigma_2} \left( \frac{\sigma}{\sigma - 1} (1 + \tau) MC_t + \mathbb{E}_t \left[ \frac{\beta f}{R} (1 - \lambda_j) \left( \frac{P_{j,t+1}}{P_{j,t}} \right)^\sigma \right] V_{j,t+1} \right] \]

### A.3 Steady State Details

In this section, I provide the fuller derivations of the steady state. The conditions are summarized in section 1.4.2

#### A.3.1 Markup and Inflation Target

Firstly, note that in steady state:

\[ \bar{\Pi} = \bar{\Pi}_j = \Pi^* \]

Equation (2.30) can be rewritten to get a steady state equation for \( \frac{P_j^*}{P_j} \) which is shown in equation (B.12). Equation (2.29) can be rewritten to get a steady state equation for \( \bar{\nu}_j \) which is shown in equation (B.13).

\[ \left( \frac{P_j^*}{P_j} \right) = \left( \frac{1 - \frac{1 - \lambda_j}{\Pi^* \lambda_j}}{\lambda_j} \right)^{\frac{1}{\sigma}} \]

\[ \bar{\nu}_j = \frac{1}{1 - (1 - \lambda_j) \Pi^* \lambda_j} \left( \frac{P_j^*}{P_j} \right)^{-\sigma} \]
In steady state equations (1.66) to (1.68) become respectively:

\[
\frac{P_j}{P} = \tilde{V}_j \left( \frac{P_j^\pi}{P_j} \right)^{-1} \tag{A.32}
\]

\[
\bar{U}_j = \frac{1}{1 - \frac{\beta_f}{R}(1 - \lambda_j)\tilde{\Pi}^{-1}} \left( \frac{P_j}{P} \right)^{-\sigma_2} \tilde{Y} \tag{A.33}
\]

\[
\bar{V}_j = \frac{1}{1 - \frac{\beta_f}{R}(1 - \lambda_j)\tilde{\Pi}^{-1}} \left( \frac{P_j}{P} \right)^{-\sigma_2} \tilde{Y} \frac{\sigma}{\sigma - 1} (1 + \tau)MC \tag{A.34}
\]

Equations (A.33) and (A.34) can be inputted into equation (A.32) to find equation (B.14).

\[
\left( \frac{P_j}{P} \right) = \frac{\sigma}{\sigma - 1} \frac{1}{1 - (1 - \lambda_j)\beta_R \tilde{\Pi}^{-1}} \left( \frac{P_j^\pi}{P_j} \right)^{-1} (1 + \tau)MC \tag{A.35}
\]

Equation (2.22) can be rewritten as equation (B.15). \( \frac{P_j}{P} \) can then be input from equation (B.14) into equation (B.15) to get equation (B.16).

\[
\int_0^1 a_j \left( \frac{P_j}{P} \right)^{1-\sigma_2} dj = 1 \tag{A.36}
\]

\[
((1 + \tau)MC)^{1-\sigma_2} \int_0^1 a_j \left[ \frac{\sigma}{\sigma - 1} \frac{1 - (1 - \lambda_j)\beta_R \tilde{\Pi}^{-1}}{1 - (1 - \lambda_j)\beta_R \tilde{\Pi}^{-1}} \left( \frac{P_j^\pi}{P_j} \right)^{-1} \right]^{1-\sigma_2} dj = 1 \tag{A.37}
\]

\( MC \) can be backed out from equation (B.16). \( \frac{\bar{P}}{\bar{P}} \) can then be found from equation (B.14). \( \bar{\nu} \) can be backed out by its definition (equation (2.27)). \( m \) can then be obtained by its definition (equation (1.57)).
A.3.2 Relative Asset Supply

The total assets supply available for the household to hold are capital and the value of firms given in equation (A.38)\(^1\)

\[ \bar{A}^s = (1 + n)\bar{K} + \bar{Z} \]  

(A.38)

Applying equation (1.26) to equation (1.29) in the steady state yields a standard equation for the value of firms equation (A.39).

\[ \bar{Z} = \frac{(1 + n)(\bar{\Omega} + \bar{Z})}{\bar{R}} \]  

(A.39)

\[ \bar{Z} = \frac{\bar{\Omega}}{\frac{\bar{R}}{1+n} - 1} \]  

(A.40)

Inputting equation (A.40) into equation (A.38) yields equation (A.41). Inputting equation (1.54) into equation (A.41) yields equation (A.42)

\[ \bar{A}^s = \frac{(\bar{R} - (1 + n))\bar{K} + \bar{\Omega}}{\frac{\bar{R}}{1+n} - 1} \]  

(A.41)

\[ \bar{A}^s = \frac{\bar{Y} - (\delta + n)\bar{K} - \bar{W}\bar{L}}{\frac{\bar{R}}{1+n} - 1} \]  

(A.42)

Equations (1.59) and (1.60) can be combined to find equation (A.43). Dividing equation (A.42) by labor income and inputting equations (1.60) and (A.43) yields equation (A.44).

\[ \bar{W}\bar{L} = \frac{1 - \alpha}{\alpha} (\bar{R} - 1 + \delta)\bar{K} \]  

(A.43)

\[ \bar{a}^s = \frac{m}{1-\alpha} - \frac{\alpha}{1-\alpha} \frac{\delta+n}{\bar{K}\bar{L}} - 1 \]  

(A.44)

\(^1\)\(\bar{K}\) needs to be multiplied by the population growth from one period to the next since assets are the assets that agents hold going forward to the next period. \(\bar{K}\) represents the per capita capital held at the start of a period. To have \(\bar{K}\) at the start of the next period, households must save \((1 + n)\bar{K}\) at the end of the previous period.
A.3.3 Relative Asset Demand

**Relative Labor** Applying arbitrage conditions on bonds and shares (equations (1.25) and (1.29)) to equations (1.40) and (1.41) yields the result that assets held at the start of $t+1$ are the assets that were saved at the end of period $t$ plus the return $\bar{R}$:

$$T = \bar{R}S^p$$

We can then apply this to equation (1.43) to yield:

$$\bar{C}_i + \bar{A}_{i+1} = \bar{W}\bar{L}_i + \bar{R}\bar{A}_i \quad (A.45)$$

where:

$$\bar{A}_i = \bar{s}_i \bar{T}$$

Next, iterate over equation (A.45) for a household from their first period of life to their last to get their intertemporal steady state budget constraint equation (A.46):

$$\sum_{i=0}^{M-1} \frac{\bar{C}_i}{\bar{R}^i} = \bar{W} \sum_{i=0}^{M-1} \frac{\bar{L}_i}{\bar{R}^i} \quad (A.46)$$

The Euler condition (equation (1.26)) can be rewritten as equation (A.47).

$$C_{i+1} = (\beta R)^{1/\gamma} C_i \quad (A.47)$$

Iterating over equation (A.47) yields equation (A.48).

$$C_i = (\beta R)^{1/\gamma} C_0 \quad (A.48)$$
Inputting this back into equation (A.46) and simplifying yields equation (A.49).

\[
\bar{C}_0 = \left( \frac{M}{\sum_{i=0}^{M-1} \beta \bar{R}_i^{\frac{(1-\gamma)}{\gamma}}} \right)^{-1} \bar{W} \sum_{i=0}^{M-1} \bar{L}_i \bar{R}_i \tag{A.49}
\]

Therefore, applying equation (A.47) to equation (A.49) yields an expression for consumption in any period in terms of \( \bar{R} \):

\[
\bar{C}_i = (\beta \bar{R})^{\frac{i}{\gamma}} \left( \frac{M}{\sum_{i=0}^{M-1} \beta \bar{R}_i^{\frac{(1-\gamma)}{\gamma}}} \right)^{-1} \bar{W} \sum_{i=0}^{M-1} \bar{L}_i \bar{R}_i \tag{A.50}
\]

Relative consumption for each cohort \( i \) is defined to be consumption by that cohort divided by labor income, as in equation (1.71):

\[
\bar{c}_i = \frac{\bar{C}_i}{\bar{W} \bar{L}}
\]

Thus, equation (A.50) can be rewritten as equation (A.51)

\[
\bar{c}_i = (\beta \bar{R})^{\frac{i}{\gamma}} \left( \frac{M}{\sum_{i=0}^{M-1} \beta \bar{R}_i^{\frac{(1-\gamma)}{\gamma}}} \right)^{-1} \sum_{i=0}^{M-1} \frac{\bar{L}_i}{\bar{R}_i} \tag{A.51}
\]

**Endogenous Labor Relative Labor Supply** In the case with exogenous labor, relative consumption for each cohort has been rewritten purely in terms of \( \bar{R} \). However, in the case with endogenous labor \( \bar{L}/L \) are endogenous so the labor part of equation (A.51) (i.e. \( \sum_{i=0}^{M-1} \frac{\bar{L}_i}{\bar{R}_i} \)) needs to be rewritten in terms of \( \bar{R} \) only.

Substituting the labor-leisure condition (equation (1.28)) into the Euler condition (equation (1.26)) yields the intertemporal labor supply condition:

\[
\frac{v'(L_{i,t})}{W_t} = \beta R_{t+1} \frac{v'(L_{i,t+1})}{W_{t+1}} \tag{A.52}
\]
Applying steady state and the disutility of working function to equation (A.52) yields:

\[ x_i \bar{L}_i^\eta = \beta \bar{R} x_{i+1} \bar{L}_{i+1}^\eta \]  \hfill (A.53)

Rewriting equation (A.54):

\[ \bar{L}_{i+1} = \left( \frac{1}{\beta \bar{R}} \right)^{\frac{1}{\eta}} \left( \frac{x_i}{x_{i+1}} \right)^{\frac{1}{\eta}} \bar{L}_i \]  \hfill (A.54)

Iterating over equation (A.54):

\[ \bar{L}_i = \left( \frac{1}{\beta \bar{R}} \right)^{\frac{i}{\eta}} \left( \frac{x_0}{x_i} \right)^{\frac{i}{\eta}} \bar{L}_0 \]  \hfill (A.55)

Next, note that the (population weighted) total labor supply is given by:

\[ \bar{L} = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \right)^{-1} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \bar{L}_i \]  \hfill (A.56)

Inputting equation (A.55) into equation (A.56) yields:

\[ \bar{L} = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \right)^{-1} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \left( \frac{1}{\beta \bar{R}} \right)^{\frac{i}{\eta}} \left( \frac{x_0}{x_i} \right)^{\frac{i}{\eta}} \bar{L}_0 \]  \hfill (A.57)

Inputting \( \bar{L}_0 \) from equation (A.57) into equation (A.55) yields the relative labor supplied by each cohort given by equation (A.58).

\[ \frac{\bar{L}_i}{\bar{L}} = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \left( \frac{1}{\beta \bar{R}} \right)^{\frac{i}{\eta}} \left( \frac{x_0}{x_i} \right)^{\frac{i}{\eta}} \right)^{-1} \left( \frac{1}{\beta \bar{R}} \right)^{\frac{i}{\eta}} \left( \frac{x_0}{x_i} \right)^{\frac{i}{\eta}} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \]  \hfill (A.58)

This has some economic intuition. When \( \bar{R} \) is higher, agents supply relatively less labor when they are old since they are already getting a high return on their savings so they don’t need to work as much.
**Relative Asset Demand**  Relative assets by cohort are defined in the same way as the definition of relative assets (equation (1.71)):

\[ \bar{a}_i = \frac{\bar{A}_i}{\bar{W}\bar{L}} \]  \hspace{1cm} (A.59)

Equation (A.60) can be rewritten in terms of relative assets and relative consumption

\[ \bar{c}_i + \bar{a}_{i+1} = \frac{\bar{L}_i}{\bar{L}} + \bar{R}\bar{a}_i \]  \hspace{1cm} (A.60)

Note that \( \bar{a}_0 = \bar{a}_M = 0 \) (since agents start with zero assets and have no need for assets when they are dead). Therefore, this yields \( M - 1 \) equations from equation (A.60) and \( M - 1 \) unknowns \( \bar{a}_1, \ldots, \bar{a}_{M-1} \). Thus, \( \bar{a}_i \) can be solved for by iterating over equation (A.60) starting from the beginning or end.

Total assets \( \bar{A} \) must equal the weighted sum of assets by cohort:

\[ \bar{A} = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \right)^{-1} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \bar{A}_i \]  \hspace{1cm} (A.61)

Next, observe that the total relative asset demand is just given by the weighted sum of the relative assets held by each cohort. This can be shown by dividing equation (A.61) by labor income and applying the definition of relative assets and relative assets by cohort (equations (1.71) and (A.59)):

\[ \bar{a}^d = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \right)^{-1} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \bar{a}_i \]

Thus, it is possible to solve for \( \bar{a}^d \) using this process.
A.4 Generalized OLG Theory

In this section, results on the steady state of a generalized life cycle model with monopolistic firms are derived. To do this, the equilibrium of relative consumption is analyzed as opposed to the equilibrium of relative assets.

Define relative consumption as:
\[ c = \frac{C}{WL} \]

By summing equation (A.51), it can be shown that the demand for relative consumption is given by:
\[ \bar{c}^d = f(\bar{R}) \]

where:
\[ f(\bar{R}) = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \right)^{-1} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} (\beta \bar{R})^{\frac{i}{2}} \left( \sum_{i=0}^{M-1} \beta^\frac{i}{2} \bar{R}^{\frac{(1-n)}{2}} \right)^{-1} \sum_{i=0}^{M-1} \frac{\bar{L}_i}{L} \frac{1}{\bar{R}^i} \] (A.62)

By adjusting the resource condition, the following supply of relative consumption can be derived:
\[ \bar{c}^s = \frac{\bar{m}}{(1-\alpha)} - \frac{\alpha(\delta + n)}{(\bar{R} - 1 + \delta)(1-\alpha)} \]

To begin, I show that the real interest rate must always be greater than the population growth.

**Theorem 7** \((\bar{R} > 1 + n \text{ Required when } m > 1)\). When \(m > 1\), require \(\bar{R} > 1 + n\) for an equilibrium to hold.

**Proof.** Firstly, note that when \(m > 1\), this guarantees that \(\Omega > 0\) since:
\[ \bar{\Omega} = \bar{Y} \left( 1 - \frac{1}{m} \right) \]

The value of shares must be non-negative in equilibrium i.e. \(Z \geq 0\). By the arbitrage
condition, it is also known that:

\[ \bar{Z} \bar{R} = (1 + n)(\bar{\Omega} + \bar{Z}) \]

\[ \bar{Z} = \frac{\bar{\Omega}}{\frac{\bar{R}}{1+n} - 1} \]

Observe that for \( \bar{Z} \geq 0 \), it must be the case that \( \bar{R} \geq 1 + n \).

Next, it is demonstrated that there must always exist an equilibrium.

**Theorem 8** (Existence of Equilibria). When \( m = 1 \), \( \bar{c}^d = \bar{c}^s \).

When \( m > 1 \), there always exists an equilibrium when \( \bar{R} > 1 + n \).

**Proof.** Note the following results:

- When \( \bar{R} = 1 + n \), \( \bar{c}^d = 1 \).
- When \( m = 1 \), \( \bar{R} = 1 + n \), \( \bar{c}^s = 1 \).
- When \( m > 1 \), \( \bar{R} = 1 + n \), \( \bar{c}^s > 1 \).
- When \( \bar{R} \to \infty \), \( \bar{c}^d \to \infty \).
- When \( \bar{R} \to \infty \), \( \bar{c}^s \to \frac{m}{1-\alpha} \).

For an equilibrium, necessary condition \( \bar{c}^s = \bar{c}^d \) must be satisfied.

When \( m = 1 \) (competitive firms): Observe that \( \bar{c}^s = \bar{c}^d \) when \( \bar{R} = 1 \). This does not automatically mean that there is a valid equilibrium at \( \bar{R} = 1 \) since we have not proved that \( \bar{Z} \geq 0 \) which is necessary for the equilibrium to seem realistic.

When \( m > 1 \) (monopolistic firms): Observe that when \( \bar{R} = 1 \), \( \bar{c}^d < \bar{c}^s \) but when \( \bar{R} \) is large, \( \bar{c}^d > \bar{c}^s \). Therefore, since \( \bar{c}^d, \bar{c}^s \) are continuous functions of \( \bar{R} \), there must be a point at which they cross and thus there exists some \( \bar{R} > 1 + n \) where \( \bar{c}^d = \bar{c}^s \).
Next, it is shown that an equilibrium must always be dynamically efficient.

**Theorem 9** ($\bar{R} \geq 1+n$ Guarantees Dynamic Efficiency). When $\bar{R} \geq 1+n$, there is dynamic efficiency. This implies that the equilibrium with monopolistic firms is always dynamically efficient.

**Proof.**

$$\bar{C} = \bar{Y} - (\delta + n)\bar{K}$$

Inputting $\bar{Y}$:

$$\bar{C} = \frac{\bar{A}\bar{K}^{\alpha}L^{1-\alpha}}{\bar{\nu}} - (\delta + n)\bar{K}$$

Differentiating:

$$\frac{d\bar{C}}{d\bar{K}} = \frac{\alpha\bar{A}\bar{K}^{\alpha-1}L^{1-\alpha}}{\bar{\nu}} - (\delta + n)$$

$$= \frac{\bar{R} - 1 + \delta}{MC\bar{\nu}} - (\delta + n)$$

$$= m(\bar{R} - 1 + \delta) - (\delta + n)$$

This is greater than 0 when $\bar{R} > 1 + n$.

\[\square\]

### A.5 Calibration Details

In this section, I provide more details on the calibration which was explained in section [1.4.5](#).

#### A.5.1 Endogenous Labor Supply

Substituting the labor-leisure condition (equation (1.28)) into the Euler condition (equation (1.26)) yields the intertemporal labor supply condition:

$$\frac{v'(L_{i,t})}{W_t} = \beta R_{t+1} \frac{v'(L_{i,t+1})}{W_{t+1}}$$

(A.63)
In steady state this becomes:

\[ x_i \bar{L}_i^\eta = \beta \bar{R} x_{i+1} \bar{L}_{i+1}^\eta \]  
(A.64)

Equation (A.64) can be rewritten as:

\[ x_{i+1} = \frac{1}{\beta R} x_i \left( \frac{\bar{L}_i}{\bar{L}_{i+1}} \right)^\eta \]  
(A.65)

Iterating over this yields:

\[ x_i = \frac{1}{(\beta R)^i x_0} \left( \frac{\bar{L}_0}{\bar{L}_i} \right)^\eta \]

Set \( x_0 = 1 \). To keep things simple, I just set \( x_i \forall i > 0 \) so that the labor supply is the same as the exogenous case parameterization when \( \beta \bar{R} = 1 \) so:

\[ x_i = \left( \frac{\bar{L}_0}{\bar{L}_i} \right)^\eta \]

where the \( \bar{L} \) ratios are the same as in the exogenous labor case.

### A.6 Results Details

In this section, I provide additional results relating to section 1.5 on what would happen if the central bank were to raise inflation from 2 to 4 percent.

The extent to which changing the degree of firm discounting i.e. \( \beta_f \) impacts the results is given in table A.1.
Table A.1: Impact of Changing Firm Discounting on the Policy Experiment Results

<table>
<thead>
<tr>
<th>Firm Additional Discount ($\beta_f$)</th>
<th>0.89</th>
<th>0.935</th>
<th>0.972</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in markup (p.p.)</td>
<td>-1.45</td>
<td>-1.07</td>
<td>-0.68</td>
<td>-0.32</td>
</tr>
<tr>
<td>Change in $\frac{\text{Cold}}{\text{Cyoung}}$ (%)</td>
<td>-7.34</td>
<td>-5.12</td>
<td>-3.10</td>
<td>-1.39</td>
</tr>
<tr>
<td>Change in r (p.p.)</td>
<td>-0.54</td>
<td>-0.38</td>
<td>-0.23</td>
<td>-0.10</td>
</tr>
<tr>
<td>Change in K (%)</td>
<td>6.43</td>
<td>4.44</td>
<td>2.66</td>
<td>1.19</td>
</tr>
</tbody>
</table>

The values of $\beta_f$ in table [A.1] were picked to correspond to: $\beta_f$ is the average discounting above the expected real rate that Jagannathan et al. (2016); Graham and Harvey (2011, 2012) actually find in their surveys; $\beta_f = 0.935$ is the average discounting from WACC which is the value that is actually used; $1/0.972 - 1$ was the average real rate paid by prime borrowers relative minus the average risk free rate between 1995 and 2007.

As the degree of discounting increases, the equilibrium real rate falls by more. This is because higher discounting means that firms care less about the future so lower the markup by more when inflation rises.

A.7 Idiosyncratic Labor Model

In this section, I present an alternative household setup where households have a similar life cycle structure but they face idiosyncratic labor shocks. This is an alternative parameterization that is discussed in section 1.5. To avoid complications, it is assumed there is no aggregate uncertainty. This assumption can be made since this model is only applied in the case of a long-run (steady state) equilibrium.

A.7.1 Households

Household ages and the population are denoted in the same manner as section 1.3.1

Since there are idiosyncratic shocks within cohorts, it is necessary to consider how individuals within a cohort will respond. For each cohort, there is a continuum of individuals
denoted $h$ between 0 and 1 for each cohort $i$.

An individual $h$ of cohort $i$ at time $t$ has a budget constraint given by equation (A.66). An agent either spends their money on consumption $C_{h,i,t}$ or saves $S_{h,i+1,t+1}$ for the next period. An agent receives direct income from working an exogenously set amount $L_{h,i,t}$ at time $t$ for real wage $W$. Savings from the previous period pay a gross return of $R$.

$$C_{h,i,t} + S_{h,i+1,t+1} \leq WL_{h,i,t} + RS_{h,i,t}$$ (A.66)

The amount that each individual works is dependent upon whether that household is employed or unemployed:

$$L_{h,i,t} = \begin{cases} L_{i,t} & \text{if employed} \\ U_{i,t} & \text{otherwise} \end{cases}$$

Whether or not the individual is employed is a Markov process.

Since there is no aggregate uncertainty, all assets must return the same. Thus, there is no need to specify exactly what assets agents hold for their savings. Instead, it can just be specified that a household $h$ in cohort $i$ at time $t$ has savings $S_{h,i,t}$ without specifying which assets they hold. Also, note that it is assumed that agents born today start with zero assets ($S_{h,0,t} = 0$).

The agent has Epstein-Zin utility which allows the effects of risk aversion and income elasticity of substitution to be separated. This means their utility is defined recursively. Their relative risk aversion is denoted by $\gamma$ and their intertemporal elasticity of substitution is denoted by $\rho$:

$$V_{i,t} = \left((1 - \beta)C_{h,i,t}^{1-\rho} + \beta\mathbb{E}_t[V_{i,t+1}^{1-\gamma}],\frac{1}{1-\gamma}]^{\frac{1}{1-\rho}} \right)$$ (A.67)

Therefore, an agent of age $k$ faces the following problem:

$$\max_{\{C_{h,i,t},S_{h,i+1,t+1}\}_{i=k}^{M-1}} \mathbb{E}_t\left[ \sum_{i=0}^{M-k-1} \beta^{i-k}V_{i,t} \right]$$ (A.68)
s.t. \( \forall i \in k, \ldots, M - 1: \)

\[
V_{i,t} = \left( (1 - \beta)C_{h,i,t}^{1-\rho} + \beta(\mathbb{E}_t[V_{i,t+1}^{1-\gamma}])^{1-\rho} \right)^{\frac{1}{1-\rho}}
\]

\[
C_{h,i,t} + S_{h,i+1,t+1} \leq WL_{h,i,t} + RS_{h,i,t}
\]

(A.69)

\[
S_{M,t+M} \geq 0
\]

As in the main model, relative assets are used. A value function approach is required due to the usage of Epstein Zin. This could be done using equations (A.68) and (A.69) but then savings will be a function of the wage which is determined on the supply side. This would require mean that it would be necessary to rerun the value function iteration each time the markup is changed. Instead, like in the non-idiosyncratic shock case, assets and consumption are written in relative terms so relative consumption and savings are defined as follows:

\[
c_{h,i,t} = \frac{C_{h,i,t}}{WL}
\]

\[
s_{h,i,t} = \frac{S_{h,i,t}}{WL}
\]

The problem can then be rewritten as\(^2\)

\[
\max_{\{c_{h,i,t+i}, s_{h,i+1,t+i+1}\}_{i=0}^{M-1}} \mathbb{E}_t\left[ \sum_{i=0}^{M-k-1} \beta^{i-k}V_{i,t} \right]
\]

(A.70)

s.t. \( \forall i \in k, \ldots, M - 1: \)

\[
V_{i,t} = \left( (1 - \beta)C_{h,i,t}^{1-\rho} + \beta(\mathbb{E}_t[V_{i,t+1}^{1-\gamma}])^{1-\rho} \right)^{\frac{1}{1-\rho}}
\]

\[
c_{h,i,t} + s_{h,i+1,t+1} \leq \frac{L_{h,i,t}}{L} + R_{h,i,t}
\]

(A.71)

\[
S_{M,t+M} \geq 0
\]

---

\(^2\)This has been simplified by cancelling a constant in the utility function.
This then comes down to a series of value function problems. The value of savings and labor income in the final period of the agent’s life is given by the utility of consuming all the remaining assets of the agent:

\[ V_{i,M-1} = ((1 - \beta)(\frac{L_{M-1}}{L} + R_{M-1})^{1-\rho})^{\frac{1}{1-\rho}} \]

Then, working backwards, the value of an agent of cohort \( i \)'s savings and labor income can be computed using the value of an agent of cohort \( i + 1 \)'s savings and labor income.

Observe that \( s_i \) can be computed for every age given the value of \( R \). Therefore, individual relative savings is effectively a function of \( R \). Consequently, the aggregate relative savings is a function of the real interest rate and the population weighted sum of individual cohort labor supplies:

\[ s(R) = \left( \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \right)^{-1} \sum_{i=0}^{M-1} \frac{1}{(1+n)^i} \int_0^1 s_{h,i}(R) dh \]

A.7.2 Other Computations

I have shown the derivation for the demand for relative assets. The supply of relative assets and the relationship between the markup and inflation are the same as in appendix A.3.

A.8 Empirics Robustness

In this section, I conduct robustness checks for the results presented in section 2.6. I discuss these results at the end of that section.

OECD Members Pre-1975 The sample is reduced to consider only consider countries that were members of the OECD before 1975.
Table A.2: Relationship between Inflation and the Real Rate OECD Original Members

<table>
<thead>
<tr>
<th>$RealRate_{i,t}^{10yr}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Inflation_{i,t-4,t}$</td>
<td>-0.142**</td>
<td>-0.169***</td>
<td>-0.546***</td>
<td>-0.909**</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.048)</td>
<td>(0.072)</td>
<td>(0.280)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>966</td>
<td>966</td>
<td>966</td>
<td>651</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Low Inflation ($< 10\%$)  The sample is reduced to consider only data points where long-run inflation exceeded 10%.

Table A.3: Relationship between Inflation and the Real Rate Low Inflation

<table>
<thead>
<tr>
<th>$RealRate_{i,t}^{10yr}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Inflation_{i,t-4,t}$</td>
<td>-0.033</td>
<td>-0.059</td>
<td>-0.643***</td>
<td>-0.948***</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.067)</td>
<td>(0.085)</td>
<td>(0.171)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1088</td>
<td>1088</td>
<td>1088</td>
<td>808</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Pre-2000  The sample is reduced to consider only before 2000.
Table A.4: Relationship between Inflation and the Real Rate Pre-2000

<table>
<thead>
<tr>
<th>$RealRate_{i,t}^{10%}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Inflation}_{i,t-4,t}$</td>
<td>-0.323***</td>
<td>-0.354***</td>
<td>-0.426***</td>
<td>-0.451</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.034)</td>
<td>(0.049)</td>
<td>(NaN)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>558</td>
<td>558</td>
<td>558</td>
<td>253</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Post-2000 The sample is reduced to consider only 2000 and after.

Table A.5: Relationship between Inflation and the Real Rate Post-2000

<table>
<thead>
<tr>
<th>$RealRate_{i,t}^{10%}$</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Inflation}_{i,t-4,t}$</td>
<td>-0.228**</td>
<td>-0.414***</td>
<td>-0.868***</td>
<td>-1.038***</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.093)</td>
<td>(0.108)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>593</td>
<td>593</td>
<td>593</td>
<td>580</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Spurious Regression Checks We consider differenced variables in order to check for spuriousness.
Table A.6: Relationship between Inflation and the Real Rate Spurious Regression Difference Check

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta RealInterest_{i,t}^{10yr}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta Inflation_{i,t-4,t}$</td>
<td>-0.611***</td>
<td>-0.702***</td>
<td>-0.850***</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.079)</td>
<td>(0.100)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>year dummies</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>N</td>
<td>1116</td>
<td>1116</td>
<td>812</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.
Appendix B

Appendix to Chapter 2
Figure B.1: Inflation and the Profit Share in the US

Profit share is measured by corporate profits before tax (without IVA and CCAdj). Inflation is CPI inflation.

B.1 Introduction Details

There is evidence of a negative relationship between inflation and firm profits in US data. We can observe a negative relationship between trend inflation and profitability in the US in figure B.1. In the 1960s and early 1970s, inflation remained subdued while profits were high. However, after inflation rose in the late 1970s, profits fell. As inflation has fallen steadily since 1990, corporate profits have risen again.
B.2 Simple Model Additional Details

B.2.1 Log Linearization of Calvo

We assume that $\bar{\Pi} = 1$ which implies that $\frac{\overline{P^*}}{P} = \frac{\sigma}{\sigma - 1} \bar{M}$. Also note that $\bar{M}_{t+j} = \beta^j$. We can then log-linearise:

$$\sum_{j=0}^{\infty} \beta^j (1 - \lambda)^j \bar{\Pi}^{j(\sigma - 1)} \frac{\overline{P^*}}{P} \left( \frac{P^*_t}{P_t} \right) = \sum_{j=0}^{\infty} \beta^j (1 - \lambda)^j \bar{\Pi}^{j(\sigma - 1)} \frac{\overline{P^*}}{P} \left[ \bar{M}_{t+j} + \left( \frac{P_{t+j}}{P_t} \right) \right]$$

Simplifying:

$$\left( \frac{P^*_t}{P_t} \right) = (1 - \beta(1 - \lambda)) \sum_{j=0}^{\infty} \beta^j (1 - \lambda)^j \left[ \bar{M}_{t+j} + \left( \frac{P_{t+j}}{P_t} \right) \right]$$

We can rewrite this iteratively as:

$$\left( \frac{P^*_t}{P_t} \right) = (1 - \beta(1 - \lambda)) \bar{M}_{t} + \beta(1 - \lambda) \left( \frac{P^*_t}{P_t} \right) \quad (B.1)$$

Next, we note that by equation (C.2):

$$P_{t+1}^{1-\sigma} = \lambda P^*_{t+1}^{1-\sigma} + (1 - \lambda) P_t^{1-\sigma}$$

Rearranging:

$$\Pi_{t+1}^{1-\sigma} = \lambda \left( \frac{P^*_t}{P_t} \right)^{1-\sigma} \Pi_{t+1}^{1-\sigma} + (1 - \lambda)$$

Log-linearisation yields:

$$\hat{\Pi}_{t+1} = \lambda \left( \frac{P^*_t}{P_t} \right) + \hat{\Pi}_{t+1}$$

---

1To simplify this log-linearisation, note that if $\sum_{i=0}^{K} M_{t+i}[A_{t+i} - B_{t+i}]$ and $\bar{A}_{t+i} = \bar{B}_{t+i}$ then $\sum_{i=0}^{K} M_{t+i}A_{t+i}A_{t+i} = \sum_{i=0}^{K} M_{t+i}B_{t+i}B_{t+i}$. 141
\[
\hat{\Pi}_{t+1} = \frac{\lambda}{1 - \lambda} \left( \frac{P^*_t}{P^*_t} \right) \tag{B.2}
\]

Inputting equation (C.7) into equation (C.6) yields:

\[
\frac{1 - \lambda}{\lambda} \hat{\Pi}_t = (1 - \beta(1 - \lambda))\hat{MC}_t + \beta(1 - \lambda)(\frac{1 - \lambda}{\lambda} + 1) \tag{B.3}
\]

Simplifying yields:

\[
\hat{\Pi}_t = \kappa \hat{MC}_t + \beta \mathbb{E}_t[\hat{\Pi}_{t+1}] \tag{B.4}
\]

where:

\[
\kappa = \frac{\lambda(1 - \beta(1 - \lambda))}{1 - \lambda}
\]

**B.2.2 Log Linearization of Rotemberg**

We can log-linearize equation (2.18) to yield:

\[
\hat{\Pi}_t = \kappa \hat{MC}_t + \beta \mathbb{E}_t[\hat{\Pi}_{t+1}]
\]

where:

\[
\kappa = \frac{\sigma \bar{MC}}{\mu}
\]

Therefore, we get the same log-linearized Phillips Curve as in the Calvo case (equation (B.4)) when \( \bar{\Pi} = 1 \) (and thus \( \bar{MC} = \frac{\sigma - 1}{\sigma} \)) and the \( \kappa \) terms equal the same value:

\[
\kappa_{rotem} = \frac{\sigma - 1}{\mu} = \frac{\lambda (1 - \beta(1 - \lambda))}{1 - \lambda} = \kappa_{calvo} \tag{B.5}
\]
B.3 Determinants of Long Run Relationship Additional Details

B.3.1 Proof of theorem 1

The result that raising inflation lowers the markup when \( \beta < 1 \) is now shown formally. First derivatives are applied:

\[
\frac{d\bar{\Omega}}{d\bar{\Pi}} = -\left[ \frac{(1 - \sigma)\bar{\Pi}^{-\sigma}}{\Pi^{1-\sigma} - (1 - \lambda)} - \frac{(1 - \sigma)\bar{\Pi}^{-\sigma}}{\Pi^{1-\sigma} - (1 - \lambda)\beta} + \frac{\sigma\bar{\Pi}^{-\sigma-1}}{\Pi^{-\sigma} - (1 - \lambda)} - \frac{\sigma\bar{\Pi}^{-\sigma-1}}{\Pi^{-\sigma} - (1 - \lambda)\beta} \right]
\]

Concentrating on the first square bracket:

\[
\frac{d\bar{\Omega}}{d\bar{\Pi}} \propto -\frac{(1 - \sigma)\bar{\Pi}^{-\sigma}(1 - \lambda)(1 - \beta)}{(\Pi^{1-\sigma} - (1 - \lambda)\beta)(\Pi^{1-\sigma} - (1 - \lambda))} - \frac{\sigma\bar{\Pi}^{-\sigma-1}(1 - \lambda)(1 - \beta)}{(\Pi^{-\sigma} - (1 - \lambda))(\Pi^{-\sigma} - (1 - \lambda)\beta)}
\]

Rearranging and remove \( (1 - \lambda)(1 - \beta) \) (which would not be possible if \( \beta = 1 \)):

\[
\frac{d\bar{\Omega}}{d\bar{\Pi}} \propto \frac{\bar{\Pi}^{-\sigma}}{(\Pi^{1-\sigma} - (1 - \lambda)\beta)(\Pi^{1-\sigma} - (1 - \lambda))} - \frac{\sigma}{\sigma - 1} \frac{\bar{\Pi}^{-\sigma-1}}{(\Pi^{-\sigma} - (1 - \lambda))(\Pi^{-\sigma} - (1 - \lambda)\beta)}
\]

Define \( \bar{m} = \frac{1}{MC^
u} \) and then substitute out terms in the second addition term using equation (A.10) and rearrange to get:

\[
\frac{d\bar{\Omega}}{d\bar{\Pi}} \propto \bar{\Pi}^{-\sigma}(\bar{\Pi}^{-\sigma} - (1 - \lambda))^2 - \bar{m}\bar{\Pi}^{-\sigma-1}(\bar{\Pi}^{1-\sigma} - (1 - \lambda))^2
\]

Simplifying:

\[
\frac{d\bar{\Omega}}{d\bar{\Pi}} \propto \bar{\Pi}^{-\sigma}(\bar{\Pi}^{-\sigma} - (1 - \lambda))^2 - \bar{m}\bar{\Pi}^{-\sigma-1}(\bar{\Pi}^{1-\sigma} - (1 - \lambda))^2
\]
\( \bar{m} \geq 1 \) otherwise firms make negative profits and \( \frac{d\bar{m}}{d\Pi} \) is decreasing in \( \bar{m} \). Therefore, \( \frac{d\bar{m}}{d\Pi} \) takes its highest possible value when \( \bar{m} = 1 \). I show that even in this case \( \frac{d\bar{m}}{d\Pi} < 0 \) and thus the markup always decreases in inflation. Under \( \bar{m} = 1 \), simplify to yield:

\[
\frac{d\bar{\chi}}{d\Pi} \bigg|_{\bar{m}=1} \propto \bar{\Pi}^{-3\sigma} + (1 - \lambda)^2 \bar{\Pi}^{-\sigma} - [\bar{\Pi}^{1-3\sigma} + \bar{\Pi}^{-\sigma-1}(1 - \lambda)^2]
\]

\[
\frac{d\bar{\chi}}{d\Pi} \bigg|_{\bar{m}=1} \propto (\bar{\Pi} - 1)((1 - \lambda)^2 - \bar{\Pi}^{-2\sigma})
\]

Then note that by equation (A.10), \( \bar{\Pi}^{-\sigma} \geq 1 - \lambda \) since this is needed to guarantee \( \bar{m} \) is positive. Thus, \( \frac{d\bar{m}}{d\Pi} < 0 \) so I have demonstrated that when average inflation rises, the average markup always falls under Calvo pricing when \( \beta < 1 \).

### B.3.2 Menu Costs Additional Figures

Figure [B.2] shows how raising inflation affects the probability of price rises and falls. Appendix [B.3.2] shows how raising inflation affects the absolute size of price rises and price falls. Figure [B.4] shows how raising inflation affects the reset price and the inaction region under a productivity of approximately \( A_i = 1 \).

### B.3.3 Degree of Monetary Non-Neutrality

A key determinant of the degree to which raising inflation lowers the markup is the degree of monetary non-neutrality. This makes sense since if firms are reluctant to change their prices then when inflation rises, they will set lower average markups. In figures [B.5] to [B.7], we observe that an increase in monetary non-neutrality through respectively a fall in the frequency of price change (\( \lambda \)), a rise in the convex cost of price adjustment (\( \mu \)) or a rise in the menu cost (\( \mu \)) means that a fall in inflation from 2% to 0% raises the profit share by more.
Figure B.2: Inflation and Frequency of Price Change Relationship

![Figure B.2: Inflation and Frequency of Price Change Relationship](image)

Figure B.3: Inflation and Size of Absolute Price Change Relationship

![Figure B.3: Inflation and Size of Absolute Price Change Relationship](image)
Figure B.4: Inflation and Inaction Region Relationship

Notes: This graph represents the policy function for a firm with productivity $A_i = 1$.

### B.4 Extended Model Details

#### B.4.1 Multi-Sector Calvo Steady State Derivation

We denote the firm’s discount factor as $M_{t,t+i}$. The firm’s problem is:

$$\max_{P^*,Y} \mathbb{E}_t \left[ \sum_{k=0}^{\infty} M_{t,t+k} (1 - \lambda_j)^k \left[ \frac{P^*_{j,t} Y_{i,j,t+k}}{P_{t+k}} - MC_{j,t+k} Y_{i,j,t+k} \right] \right]$$

(B.6)

s.t.

$$Y_{i,j,t+k} = \left( \frac{P^*_{j,t}}{P_{j,t+k}} \right)^{-\sigma} Y_{j,t+k}$$

(B.7)

Equation (B.7) can be input into equation (B.8) to get:

$$\max_{P^*_{j,t}} \mathbb{E}_t \left[ \sum_{k=0}^{\infty} M_{t,t+k} (1 - \lambda_j)^k \left[ P^*_{j,t} \left( 1 - \sigma_{j,t+k} \right) Y_{j,t+k} - MC_{j,t+k} P^*_{j,t} \left( \frac{1}{P_{t+k}} \right) Y_{j,t+k} \right] \right]$$

(B.8)
Taking FOCs:

\[
\mathbb{E}_t \left[ \sum_{k=0}^{\infty} M_{t,t+k} (1 - \lambda_j)^k [(1 - \sigma) P_{j,t}^{* - \sigma} P_{j,t+k}^\sigma - \sigma MC_{j,t+k} + \sigma P_{j,t}^\sigma P_{j,t}^{* - \sigma - 1} Y_{j,t+k} + \sigma MC_{j,t+k} P_{j,t}^{* - \sigma - 1} P_{j,t+k} Y_{j,t+k}] \right] \tag{B.9}
\]

Rearranging:

\[
\mathbb{E}_t \left[ \sum_{k=0}^{\infty} M_{t,t+k} (1 - \lambda_j)^k P_{j,t+k}^\sigma Y_{j,t+k} \left[ \frac{P_{j,t}^*}{P_{t+k}} - \frac{\sigma}{\sigma - 1} MC_{j,t+k} \right] \right] \tag{B.10}
\]

Inputting \( Y_{j,t+k} \) and dividing by \( P_{j,t}^\sigma \):

\[
\mathbb{E}_t \left[ \sum_{k=0}^{\infty} M_{t,t+k} (1 - \lambda_j)^k \left( \frac{P_{j,t+k}}{P_{j,t}} \right)^\sigma \left( \frac{P_{j,t+k}}{P_{t+k}} \right)^{-\sigma} Y_{t+k} \left[ \frac{P_{j,t}^*}{P_{j,t}} \frac{P_{j,t}}{P_t} \frac{P_{t+k}}{P_{t+k}} - \frac{\sigma}{\sigma - 1} MC_{j,t+k} \right] \right] \tag{B.11}
\]

Figure B.5: Change in Profit Share when Inflation Falls from 2% to 0% Under Different Frequencies of Price Change in Calvo
Figure B.6: Change in Profit Share when Inflation Falls from 2% to 0% Under Different Extent of Rotemberg Costs

Now, let’s consider the steady state. Firstly, note that in steady state:

$$\bar{\Pi} = \bar{\Pi}_j = \Pi^*$$

Equation (2.30) can be rewritten to get a steady state equation for $\frac{P^*}{P_j}$ which is shown in equation (B.12). Equation (2.29) can be rewritten to get a steady state equation for $\bar{\nu}_j$ which is shown in equation (B.13).

$$\left(\frac{P_j^*}{P_j}\right)^{1-\frac{\lambda_j}{1-\frac{1-\lambda_j}{\Pi^*}}^{\frac{1}{1-\sigma}}}$$

(B.12)

$$\bar{\nu}_j = \frac{1}{1 - (1 - \lambda_j)\Pi^* \lambda_j \left(\frac{P_j^*}{P_j}\right)^{-\sigma}}$$

(B.13)
In steady state, equation (B.11) becomes:

$$\left( \frac{P_j}{\bar{P}} \right) = \frac{\sigma}{\sigma - 1} \frac{1 - (1 - \lambda_j) \beta_j \Pi^\sigma - 1 \left( \frac{P_j^*}{P_j} \right)}{1 - (1 - \lambda_j) \beta_j \Pi^\sigma - 1 \left( \frac{P_j^*}{P_j} \right)} - 1 \mathrm{MC}$$  \hspace{1cm} (B.14)

Equation (2.22) can be rewritten as equation (B.15). \( \left( \frac{P_j}{\bar{P}} \right) \) can then be input from equation (B.14) into equation (B.15) to get equation (B.16).

$$\int_0^1 a_j \frac{(P_j)}{\bar{P}}^{1-\sigma_2} dj = 1$$  \hspace{1cm} (B.15)

$$\overline{MC}^{1-\sigma_2} \int_0^1 a_j \left[ \frac{\sigma}{\sigma - 1} \frac{1 - (1 - \lambda_j) \beta_j \Pi^\sigma - 1 \left( \frac{P_j^*}{P_j} \right)}{1 - (1 - \lambda_j) \beta_j \Pi^\sigma - 1 \left( \frac{P_j^*}{P_j} \right)} \right]^{1-\sigma_2} dj = 1$$  \hspace{1cm} (B.16)

\( \overline{MC} \) can be backed out from equation (B.16). \( \left( \frac{P_j}{\bar{P}} \right) \) can then be found from equation (B.14). \( \bar{\nu} \) can be backed out by its definition (equation (2.27)).
B.5 Empirics Robustness

In this section, I conduct robustness checks for the results presented in section 2.6. I discuss these results at the end of that section.

OECD Members Pre-1975 The sample is reduced to consider only consider countries that were members of the OECD before 1975.

Table B.1: Relationship between Inflation and the Labor Share OECD Original Members

<table>
<thead>
<tr>
<th>LaborShare_{i,t}</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation_{i,t-4,t}</td>
<td>0.269</td>
<td>0.517*</td>
<td>0.164</td>
<td>0.351</td>
</tr>
<tr>
<td></td>
<td>(0.406)</td>
<td>(0.235)</td>
<td>(0.441)</td>
<td>(0.974)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>N</td>
<td>584</td>
<td>584</td>
<td>584</td>
<td>500</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Low Inflation ($< 10\%$) The sample is reduced to consider only data points where long-run inflation exceeded 10%.

Table B.2: Relationship between Inflation and the Labor Share Low Inflation

<table>
<thead>
<tr>
<th>LaborShare_{i,t}</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation_{i,t-4,t}</td>
<td>-0.061</td>
<td>0.719**</td>
<td>0.753*</td>
<td>0.514</td>
</tr>
<tr>
<td></td>
<td>(0.512)</td>
<td>(0.242)</td>
<td>(0.312)</td>
<td>(0.416)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>N</td>
<td>745</td>
<td>745</td>
<td>745</td>
<td>678</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Pre-2000 The sample is reduced to consider only before 2000.
Table B.3: Relationship between Inflation and the Labor Share Pre-2000

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LaborShare_{i,t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Inflation_{i,t-4,t}$</td>
<td>0.180*</td>
<td>0.211***</td>
<td>0.200***</td>
<td>0.074</td>
</tr>
<tr>
<td></td>
<td>(0.080)</td>
<td>(0.043)</td>
<td>(0.048)</td>
<td>(NaN)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>286</td>
<td>286</td>
<td>286</td>
<td>198</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Post-2000  The sample is reduced to consider only 2000 and after.

Table B.4: Relationship between Inflation and the Labor Share Post-2000

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LaborShare_{i,t}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Inflation_{i,t-4,t}$</td>
<td>-0.886</td>
<td>0.867***</td>
<td>1.049***</td>
<td>0.856*</td>
</tr>
<tr>
<td></td>
<td>(0.804)</td>
<td>(0.254)</td>
<td>(0.282)</td>
<td>(0.339)</td>
</tr>
<tr>
<td>country dummies</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>year dummies</td>
<td></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>527</td>
<td>527</td>
<td>527</td>
<td>523</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.

Spurious Regression Checks  We consider differenced variables in order to check for spuriousness.
Table B.5: Relationship between Inflation and the Labor Share Spurious Regression Difference Check

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \text{LaborShare}_{i,t}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \text{Inflation}_{i,t-4,t}$</td>
<td>0.189***</td>
<td>0.187***</td>
<td>0.177***</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.049)</td>
<td>(0.052)</td>
</tr>
<tr>
<td>year dummies</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>econ controls</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>N</td>
<td>781</td>
<td>781</td>
<td>707</td>
</tr>
</tbody>
</table>

Notes: *, ** and *** respectively represent $p < 0.05$, $p < 0.01$ and $p < 0.001$. The numbers in parentheses represent the clustered standard errors.
Appendix C

Appendix to Chapter 3
C.1 Model Setup Details

C.1.1 Firm Details

CES Aggregator Firm

We have a competitive aggregator firm which produces final goods from a continuum of intermediate goods firms. It has the following problem:

\[
\min_{[Y_i, t]} \int_0^1 P_{i,t}Q_{i,t} di
\]

s.t.

\[
Y_t = \left( \int_0^1 Y_{i,t}^{\sigma-1} \right)^{\frac{\sigma}{\sigma-1}}
\]

We get the following first order conditions:

\[
Y_{i,t} = Y_t \left( \frac{P_{i,t}}{P_t} \right)^{-\sigma} \tag{C.1}
\]

\[
P_t = \left( \int_0^1 P_{i,t}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{C.2}
\]

Intermediate Goods Production

Intermediate goods firms have a linear production function over labor:

\[
Y_{i,t} = A_t L_{i,t}
\]

Inputting equation (C.1) and aggregating yields:

\[
Y_t \int_0^1 \left( \frac{P_{i,t}}{P_t} \right)^{-\sigma} di = A_t L_t
\]
Log-linearising and noting that \( \int_0^1 \left( \frac{P_i}{P_t} \right)^{-\sigma} \, di = 0 \) yields:

\[
\hat{Y}_t = \hat{A}_t + \hat{L}_t
\] (C.3)

They therefore face the following constant marginal cost to produce:

\[
MC_t = \frac{W_t}{A_t}
\] (C.4)

Log linearising equation (C.10) yields:

\[
\hat{MC}_t = \hat{W}_t - \hat{A}_t
\] (C.5)

**Intermediate Goods Pricing**

Each intermediate goods firm is monopolistic. It sets a price \( P_{i,t} \) taking into account the demand for its good from the final goods firm. It only changes its price with probability \( \lambda \). It discounts a period \( t + i \) periods in the future by \( M_{t,t+i} \). Here is its problem in real terms:

\[
\max_{P_{i,t}} \mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j \left[ \frac{P_{i,t} Y_{i,t+j}}{P_{t+k}} - MC_{t+j} Y_{i,t+j} \right] \right]
\]

s.t.

\[
Y_{i,t} = Y_t \left( \frac{P_{i,t}}{P_t} \right)^{-\sigma}
\]

Inputting demand:

\[
\max_{P_{i,t}} \mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j Y_{i,t+j} \left( \frac{P_{t+j} P_{i,t}}{P_t} \right)^{1-\sigma} \left[ \left( \frac{P_{i,t}}{P_t} \right)^{1-\sigma} - MC_{t+j} \frac{P_{t+j}}{P_t} \left( \frac{P_{i,t}}{P_t} \right)^{-\sigma} \right] \right]
\]
Taking FOCs:

\[
\mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j Y_{t+j} \left( \frac{P_{t+k}}{P_t} \right)^{\sigma-1} \left[ (1 - \sigma) \left( \frac{P_t^*}{P_t} \right)^{-\sigma} + \sigma M_{C,t+j} \frac{P_{t+j}}{P_t} \left( \frac{P_t^*}{P_t} \right)^{-\sigma-1} \right] \right] = 0
\]

Simplifying:

\[
\mathbb{E}_t \left[ \sum_{j=0}^{\infty} M_{t,t+j} (1 - \lambda)^j Y_{t+j} \left( \frac{P_{t+k}}{P_t} \right)^{\sigma-1} \left[ \frac{P_t^*}{P_t} - \frac{\sigma}{\sigma-1} M_{C,t+j} \frac{P_{t+j}}{P_t} \right] \right] = 0
\]

We assume that \( \bar{\Pi} = 1 \) which implies that \( \frac{P_t^*}{P_t} = \frac{\sigma}{\sigma-1} \bar{M}C \). Also note that \( \bar{M}_{t,t+j} = \beta^j \).

We can then log-linearise:\(^1\)

\[
\sum_{j=0}^{\infty} \beta^j (1 - \lambda)^j \bar{Y} \bar{\Pi}^{(\sigma-1)} \frac{P_t^*}{P_t} = \sum_{j=0}^{\infty} \beta^j (1 - \lambda)^j \bar{Y} \bar{\Pi}^{(\sigma-1)} \frac{P_t^*}{P_t} \left[ \bar{M}_{C,t+j} + \left( \frac{P_{t+j}}{P_t} \right) \right]
\]

Simplifying:

\[
\left( \frac{P_t^*}{P_t} \right) = (1 - \beta (1 - \lambda)) \sum_{j=0}^{\infty} \beta^j (1 - \lambda)^j \left[ \bar{M}_{C,t+j} + \left( \frac{P_{t+j}}{P_t} \right) \right]
\]

We can rewrite this iteratively as:

\[
\left( \frac{P_t^*}{P_t} \right) = (1 - \beta (1 - \lambda)) \bar{M}C_t + \beta (1 - \lambda) \left( \frac{P_{t+1}^*}{P_t} \right) \tag{C.6}
\]

Next, we note that by equation \( \text{(C.2)} \):

\[
P_{t+1}^{1-\sigma} = \lambda P_{t+1}^{1-\sigma} + (1 - \lambda) P_t^{1-\sigma}
\]

Rearranging:

\[
\Pi_{t+1}^{1-\sigma} = \lambda \left( \frac{P_{t+1}^*}{P_{t+1}} \right)^{1-\sigma} \Pi_{t+1}^{1-\sigma} + (1 - \lambda)
\]

\(^1\)To simplify this log-linearisation, note that if \( \sum_{i=0}^{K} M_{t+i} [A_{t+i} - B_{t+i}] \) and \( \bar{A}_{t+i} = \bar{B}_{t+i} \) then \( \sum_{i=0}^{K} M_{t+i} \bar{A}_{t+i} \bar{A}_{t+i} = \sum_{i=0}^{K} M_{t+i} \bar{B}_{t+i} \bar{B}_{t+i} \).

156
Log-linearisation yields:

\[
\hat{\Pi}_{t+1} = \lambda \left( \frac{P^*_{t+1}}{P_{t+1}} + \hat{\Pi}_{t+1} \right)
\]

\[
\hat{\Pi}_{t+1} = \frac{\lambda}{1 - \lambda} \left( \frac{P^*_{t+1}}{P_{t+1}} \right) \tag{C.7}
\]

Inputting equation (C.7) into equation (C.6) yields:

\[
\frac{1 - \lambda}{\lambda} \hat{\Pi}_t = (1 - \beta(1 - \lambda))\hat{MC}_t + \beta(1 - \lambda)\left( \frac{1 - \lambda}{\lambda} + 1 \right) \tag{C.8}
\]

Simplifying yields:

\[
\hat{\Pi}_t = \kappa_2 \hat{MC}_t + \beta \mathbb{E}_t[\hat{\Pi}_{t+1}] \tag{C.9}
\]

where:

\[
\kappa_2 = \frac{\lambda(1 - \beta(1 - \lambda))}{1 - \lambda}
\]

**Substituting out the Marginal Cost**

Inputting equation (3.12) into equation (C.10) yields:

\[
\hat{MC}_t = \frac{1}{\gamma} \hat{C}_t + \eta \hat{L}_t - \hat{A}_t \tag{C.10}
\]

Inputting equation (C.3) into equation (C.10) and noting that \( \hat{C}_t = \hat{Y}_t \) yields:

\[
\hat{MC}_t = (\frac{1}{\gamma} + \eta)\hat{Y}_t - (1 + \eta)\hat{A}_t \tag{C.11}
\]

In the case where there are no price rigidities, which we denote with a superscript \( n \), intermediate goods firms will always set their marginal cost to equal \( \frac{\sigma - 1}{\sigma} \). Therefore, without price rigidities, equation (C.11) becomes:

\[
0 = (\frac{1}{\gamma} + \eta)\hat{Y}_t^n - (1 + \eta)\hat{A}_t \tag{C.12}
\]
Subtracting equation (C.12) from equation (C.11) and recalling that \( x_t = \hat{Y}_t - \hat{Y}_t^n \) yields:

\[
\hat{MC}_t = \left( \frac{1}{\gamma} + \eta \right) x_t \tag{C.13}
\]

Inputting equation (C.13) into equation (C.9) yields:

\[
\hat{\Pi}_t = \kappa x_t + \beta E_t[\hat{\Pi}_{t+1}]
\]

where:

\[
\kappa = \kappa_2 \left( \frac{1}{\gamma} + \eta \right)
\]

### C.2 Determinacy

#### C.2.1 Proof of Theorem 5

To consider stability/determinacy, we need to analyze the full system of equations i.e. equations (3.13), (3.15) and (3.33). We simplify by inputting equation (3.15) into equation (3.33). We then rearrange to get:

\[
E_t[Z_{t+1}] = AZ_t
\]

where \( Z_t = (x_t, \pi_t)' \):

\[
A = \begin{pmatrix}
\frac{1}{M} (1 + \frac{\sigma \kappa}{\beta}) & \frac{\sigma}{M} (\phi_\pi - \frac{1}{\beta}) \\
-\frac{\kappa}{\beta} & \frac{1}{\beta}
\end{pmatrix}
\]

\( x_{t+1}, \pi_{t+1} \) are both controls. Therefore, for determinacy and stability, we require that the eigenvalues of \( A \) are greater than 1.

We find the characteristic equation \( P(\lambda) = |A - \lambda I| \):

\[
P(\lambda) = \left( \frac{1}{M} \left( 1 + \frac{\sigma \kappa}{\beta} \right) - \lambda \right) \left( \frac{1}{\beta} - \lambda \right) - \frac{\sigma}{M} \left( \phi_\pi - \frac{1}{\beta} \right) \left( -\frac{\kappa}{\beta} \right)
\]
\[ P(\lambda) = \lambda^2 - \left( \frac{1}{M} \left( 1 + \frac{\sigma \kappa}{\beta} \right) + \frac{1}{\beta} \right) \lambda + \frac{1}{M \beta} + \frac{\sigma \phi \pi \kappa}{M \beta} \]

When \( \lambda \) is very negative or positive, the first \( \lambda^2 \) term dominates. The minimum of \( P \) can be found by taking the FOC of \( P \) and setting it to zero:

\[ \lambda_{\text{min}} = 0.5 \left( \frac{1}{M} \left( 1 + \frac{\sigma \kappa}{\beta} \right) + \frac{1}{\beta} \right) \]

We see that \( \lambda_{\text{min}} > 1 \). We also observe that if \( P \) has complex roots \( a + \sqrt{b}i, a - \sqrt{b}i \) then \( P(\lambda) = (\lambda - a - \sqrt{b}i)(\lambda - a + \sqrt{b}i) = (\lambda - a)^2 + b \). We observe that \( P \) is minimized when \( \lambda = a \) so \( a = \lambda_{\text{min}} \). Therefore, in the case where we have complex roots, the system is determinate.

Since \( P \) is falling at \( \lambda = 1 \) but increases above \( \lambda_{\text{min}} \), this implies that both eigenvalues are greater than one iff \( P(1) > 0 \). Therefore, we have determinacy iff \( P(1) > 0 \) which is shown in equation (C.14). When \( M = 1 \), we require \( \phi_\pi > 1 \) to get \( P(1) > 0 \). However, when we allow \( M < 1 \), we no longer require \( \phi_\pi > 1 \) to get \( P(1) > 0 \) and we can even get \( P(1) > 0 \) when \( \phi_\pi = 0 \) (the fixed interest rate case):

\[ P(1) = \left( 1 - \frac{1}{M} \right) \left( 1 - \frac{1}{\beta} \right) + \frac{\sigma \kappa}{M \beta} (\phi_\pi - 1) \]  

\[ \text{(C.14)} \]

C.2.2 Proof of Theorem 6

To consider stability/determinacy, we need to analyze the full system of equations i.e. equations (3.13), (3.15), (3.40) and (3.41). We simplify this system by inputting equation (3.15) into equation (3.41). We find:

\[ \mathbb{E}_t[Z_{t+1}] = AZ_t \]
where:

\[
Z_t = \begin{pmatrix} x_t \\ \pi_t \\ \mu_{x,t} \end{pmatrix}
\]

\[
A = \begin{pmatrix}
\frac{1}{\bar{M}}(1 + \sigma \frac{\kappa}{\beta}) & \frac{1}{\bar{M}}\sigma(\phi_\pi - \frac{1}{\beta}) & -\frac{1}{\bar{M}}(1 - \bar{M}) \\
-\frac{\kappa}{\beta} & \frac{1}{\beta} & 0 \\
\alpha_y & 0 & 1 - \alpha_y
\end{pmatrix}
\]

\(x_t, \pi_t\) are controls while \(\mu_{x,t}\) is a state. Let’s denote the three eigenvalues by \(\lambda_1, \lambda_2, \lambda_3\) and \(r(\lambda)\) represents the real component of \(\lambda\). To get a unique, stable solution we require that \(|r(\lambda_1)| < 1, |r(\lambda_2)|, |r(\lambda_3)| > 1\). If \(|r(\lambda_1)|, |r(\lambda_2)|, |r(\lambda_3)| \geq 1\) then we have instability. If \(|r(\lambda_1)|, |r(\lambda_2)| \leq 1\) then we have indeterminacy.

We have the following characteristic equation:

\[
P(\lambda) = \lambda^3 - \left(\frac{1}{\bar{M}}(1 + \sigma \frac{\kappa}{\beta}) + \frac{1}{\beta} + 1 - \alpha_y\right)\lambda^2
+ \left[\frac{1}{\bar{M}}(1 + \sigma \frac{\kappa}{\beta})(\frac{1}{\beta} + 1 - \alpha_y) + \frac{1}{\beta}(1 - \alpha_y) + \frac{1}{\bar{M}}\sigma(\phi_\pi - \frac{1}{\beta})(\frac{\kappa}{\beta}) + (\frac{1}{\bar{M}}(1 - \bar{M}))\alpha_y\right]\lambda
- \left[\frac{1}{\bar{M}}(1 + \sigma \frac{\kappa}{\beta})\frac{1}{\beta}(1 - \alpha_y) + \frac{1}{\bar{M}}\sigma(\phi_\pi - \frac{1}{\beta})(\frac{\kappa}{\beta})(1 - \alpha_y) + (\frac{1}{\bar{M}}(1 - \bar{M}))\frac{1}{\beta}\alpha_y\right]
\]

\[(C.15)\]

**Lemma 10** \((P(1) > 0 \text{ is a Necessary Condition for Determinacy})\). \(P\) is a cubic polynomial with a positive coefficient on the cubic term: If \(P(1) \leq 0\) then there cannot exist one root of the polynomial with real part strictly less than 1 and two roots with real part strictly greater than 1.

**Proof.** We proceed a proof by contradiction. We consider three cases.

Firstly, if \(P(1) = 0\) then 1 is a root and we require roots strictly less/greater than 1 so this cannot satisfy the stated condition.

Secondly, we consider the case where all roots are real. In this case, \(P(\lambda)\) crosses zero
three times. We denote these $\lambda_1 \leq \lambda_2 \leq \lambda_3$. Since the coefficient on the cubic term in the polynomial is positive, when $\lambda < \lambda_1$, $P(\lambda) < 0$ and when $\lambda > \lambda_3$, $P(\lambda) > 0$. Therefore, at points satisfying $\lambda_1 < \lambda < \lambda_2$, $P(\lambda) > 0$ and at points satisfying $\lambda_2 < \lambda_3$, $P(\lambda) < 0$. Therefore, we see that if $P(1) < 0$ we must either have at $1 < \lambda_1$ so all roots are greater than 1 or $\lambda_2 < 1 < \lambda_3$ so exactly one root is greater than 1. Therefore, it is not possible that with three real roots and $P(1) < 0$ that we have exactly one root which is strictly less than 1.

Thirdly, if $P$ has complex roots then it must have exactly two (since they come in pairs) i.e. $\lambda_2 = b + ci, \lambda_3 = b - ci$. We can then write:

$$P(\lambda) = (\lambda - \lambda_1)(\lambda - b - ci)(\lambda - b + ci)$$

We see from equation (C.16) that if $P(1) < 0$ then $a > 1$ (since the second bracket is strictly positive). In this case, we must have that either $b < 1$ in which case there are exactly two roots with real parts strictly less than 1 or $b > 1$ in which case there are exactly zero roots with real parts strictly less than 1. Therefore, we see that if $P(1) < 0$ and we have complex roots that we cannot have exactly one root strictly less than 1.

Equation (C.17) shows the characteristic equation evaluated at 1. We see that $P(1) > 0$ only when $\phi_\pi > 1$. Therefore, by lemma [10], the only way it is possible to get exactly one root with real part strictly less than 1 and two roots with real part strictly greater than 1 is if $\phi_\pi > 1$. Consequently, we see that whenever $\phi_\pi \leq 1$, we must either have indeterminacy or instability.

$$P(1) = \frac{\alpha_y}{M_\beta \kappa \sigma} (\phi_\pi - 1) \quad (C.17)$$

\[\text{2}^2\] Under $\phi_\pi \leq 1$, there are parameter values for which we get indeterminacy and there are parameter values for which we get stability.
C.3 Costs of Zero Lower Bound

C.3.1 Costs of Zero Lower Bound in Standard Behavioral New Keynesian Model

Call $r_z$ the low level of $r$ that the economy gets stuck at indefinitely. We must have that:

$$r_z < -\pi^*$$

Then, denoting $\hat{i}_z$ to be the value that $\hat{i}_z$ takes at the zero lower bound:

$$\hat{r}_z \leq -\pi^* - r^n = -\bar{r} = \hat{i}_z$$

We set $\hat{r}_z - \hat{i}_z = -0.01$. We can then behavioral IS curve as:

$$\hat{x}_t = M\mathbb{E}_t[\hat{x}_{t+1}] - \sigma(0.01 - \mathbb{E}_t[\hat{r}_{t+1}])$$  \hspace{1cm} (C.18)

The NKPC stays the same.

We also note that when the ZLB does end and we return to normal then if we follow a simple Taylor rule i.e. $i_t = \phi \pi_t$ then we would need to have $\pi_t = 0$ at the conclusion of the ZLB and this would require $x_t = 0$. We can then just work backwards from this final period to simulate what will happen.

Another way of observing this is to observe that $\hat{x}$ can converge to a steady state level. The steady state of the IS and NKPC curves is respectively:

$$\bar{x} = M\bar{x} - \sigma(0.01 - \bar{\pi})$$

$$\bar{\pi} = \kappa \bar{x} + \beta \bar{\pi}$$

162
Combining these:

\[ \tilde{x} = M \tilde{x} - \sigma (0.01 - \frac{\kappa}{1 - \beta} \tilde{x}) \]

Rearranging:

\[ \left[ 1 - M - \frac{\kappa}{1 - \beta} \right] \tilde{x} = -\sigma (0.01) \]

As long as the term in square brackets is positive, \( \tilde{x} \) will converge.

### C.3.2 Costs of Zero Lower Bound in Long-Term Behavioral New Keynesian Model

We provide a sketch proof.

If \( \bar{\mu}_{\tilde{x}} = 0 \) then:

\[ \bar{\mu}_{\tilde{x}} = M \tilde{x} - \sigma (0.01 - \frac{\kappa}{1 - \beta} \tilde{x}) \] (C.19)

Rearranging:

\[ \left[ 1 - M - \frac{\kappa}{1 - \beta} \right] \tilde{x} = -\sigma (0.01) \]

However, this implies that the long-term expectatons of \( \tilde{x} \) should actually be:

\[ \bar{\mu}_{\tilde{x}} = -\frac{1}{1 - M - \frac{\kappa}{1 - \beta}} \sigma (0.01) \]

And if we input this into equation (C.19) then we'll get that \( \tilde{x} \) is lower and thus \( \bar{\mu}_{\tilde{x}} \) falls.

And this continues indefinitely.