Cardiff University

Doctoral Thesis

Stochastic Equilibrium, the Phillips Curve and Keynesian Economics

Author:
David Staines

Supervisor:
Professor Huw Dixon

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Economics Section
Cardiff Business School

January 28, 2019

Abstract

I uncover serious problems with the benchmark New Keynesian Phillips curve linearized around its non-stochastic zero inflation steady state when the underlying model features a subset of prices that stay rigid over multiple periods, as in the popular Calvo model. I am able to demonstrate that the dynamics of approximations taken at the non-stochastic steady state are non-hyperbolic. This means that approximations taken at this point do not represent a valid description of the dynamics of the underlying model at any other point in the state space. This allows me to overturn results such as the 'Divine Coincidence' that equates welfare under price rigidity with the level prevailing under price dispersion. I introduce a dynamic stochastic concept of equilibrium that can be applied to New Keynesian models and offers a natural point to take approximations to analyze business cycle dynamics. It is methodologically interesting as it is a notion of general equilibrium that does not correspond to partial equilibrium.

Keywords: Macroeconomics, Mathematical Economics, Random Dynamical Systems, General Equilibrium, Monetary Policy

JEL Classification: C6, D5, E1, E3, E5

2010 Mathematics Subject Classification: 37Axx, 37Bxx, 37Cxx, 37Dxx, 37Gxx, 37Hxx.

Declaration of Authorship

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed .... (candidate) Date 28/01/2019

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed .... (candidate) Date 28/01/2019

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed .... (candidate) Date 28/01/2019

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate) Date 28/01/2019 ....

STATEMENT 4

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate) Date 28/01/2019 ....

Acknowledgements

I would like to thank the Economic and Social Research for offering me a scholarship that was essential to fund my research. Support staff, such as, Elsie Phillips and Lainey Clayton who have borne some of my foibles for the last five years and helped me through the PhD scheme. I would like to thank Professor Huw Dixon for his encouragement and insightfulness. Patrick Minford for helpful comments and imparting some of his passion both for DSGE modeling and macroeconomic debate. All of the staff in the Economics Section at Cardiff have given me excellent support and advice. That has helped me grow as an economist.

Furthermore, I have benefited a great deal from discussion and debate with other economists Michael Hatcher, Ceri Davis, Lorant Kaszab, Tom Wilkinson, Ashley Nanton, Antony Savagar, Tamas Csabafi, Craig Thamotheram, Tim Jackson, Tian Maoshan, George Melios, Nikos Tzivanakis and Dandan Wang come first to mind. I would also like to thank my friends in other sections of the business school and across the university who have enriched my life greatly during my time in Cardiff.

Finally, I would like to thank my examiners for comments and corrections that greatly improved and extended the analysis. Any remaining errors are my own.

Dedicated to my parents, my grandmother and my closest friends

Contents

-

Chapter 1

Introduction

The macroeconomic profession has broadly settled on a *New Neo-Classical Synthesis*, Goodfriend and King [1997]. This approach adds imperfect competition and staggered price adjustment to the skeleton of the Real Business Cycle model to generate micro-founded compliments to the aggregate demand and aggregate supply equations of Old Keynesian economics. The Euler equation specifying optimal consumption and the optimal price-setting relation - the New Keynesian Phillips curve (NKPC) - simplify to their Old Keynesian counterparts when non-contemporaneous exogenous variables are held fixed. When this constraint is relaxed, it is necessary to specify an interest rate rule consistent with the "Taylor Principle" to generate stable solution paths.[1] Thus a three equation set up emerges.

The combination of rational expectations with monetary non-neutrality allows one to analyze systematic monetary policy seemingly unhindered by the Lucas Critique, Lucas Jr [1976]. For this reason, the three equation New Keynesian framework underpins the field of Dynamic Stochastic Equilibrium (DSGE) modeling , now popular in academia and widely used by Central Banks.

Nevertheless, New Keynesianism has had problems distinguishing itself from the Neo-Classical tradition in terms of policy prescriptions and model predictions. Macroeconomics as a separate intellectual discipline came into existence following *The General Theory of Employment, Interest and Money* (Keynes [1936]), in order to learn how best to mitigate inefficient business cycle fluctuations. The lack of inefficient

[1] The "Taylor Principle" states that the real interest rate should be expected to increase in response to higher inflation in order to drive it back to target following a shock to expectations to prevent indeterminacy in the linear system associated with sunspot equilibria, see Woodford [2001]. The term was coined in respect of the seminal work on monetary policy rules Taylor [1993a]. Simultaneously, Henderson and McKibbin [1993] proposed a similar rule, which I will discuss later.

fluctuations in benchmark Real Business Cycle models, such as Long Jr and Plosser [1983] and Barro and King [1984], is the most unappealing aspect of Neo-Classical modeling and the reason why it has never enjoyed favor in policy circles.

Unfortunately, comparable New Keynesian models also suffer from this problem. Correia et al. [2008] show that the Central Bank can implement the social optimum when the government is using a standard set of distortionary tax instruments to correct static market failures. Woodford [2000], Michael [2002] show an optimal monetary policy can successfully stabilize inflation and the output gap simultaneously under a wide variety of shock processes. This result has been labeled the 'Divine Coincidence' Blanchard and Galí [2007](henceforth DC).[2] It is an anathema in policy circles. 'Inflation nutters' is the uncharitable description former Bank of England Governor Mervyn King gave to those advocating complete inflation stabilization as a policy objective, as if the DC applied King [1997].

This reflects a fundamental disjuncture between optimal monetary policy in theory and successful policy practice. I show that in the stochastic compliment to the DC framework, deviations of inflation and the output gap from target should be white noise. Therefore in the limit as Central Banks become better able to observe shocks in real time and change policy rates more frequently the Central Bank can 'fine-tune'[3]away all fluctuations in inflation and the output gap.

This gives rise to an 'inflation persistence puzzle': in the data, inflation appears persistent across all nations, time periods, policy regimes, levels of aggregation and plausible assumptions about trends in other macroeconomic variables O'Reilly and Whelan [2005], Pivetta and Reis [2007], Gerlach and Tillmann [2012], Imbs et al. [2011], Meller and Nautz [2012], Kouretas and Wohar [2012], Plakandaras et al. [2014], Vaona and Ascari [2012], Tillmann [2013], Choi and O'Sullivan [2013], Nakamura

[2]The possibility of a binding zero bound on nominal interest rate overturns this conclusion, in particular it is optimal for the Central Bank to allow a period of above target inflation and output immediately following a zero bound spell Eggertsson et al. [2003]. However, the problem may reoccur if we adopt the more empirically credible assumption that Central Banks can mimic negative nominal interest rates through quantitative easing- see Debortoli et al. [2018].

[3]The term is frequently attributed to Walter Heller Chief Economic Adviser to President Kennedy see for example http://connection.ebscohost.com/c/reference-entries/40422478/fine-tuning-1960s-economics. It referred originally to fiscal policy in an 'Old Keynesian' set up. Scepticism about the concept was focal to monetarist opposition to traditional Keynesian macroeconomicsSnowdon and Vane [2005] see for example Friedman [1968b].

and Steinsson [2013].[4] The forward-looking NKPC is strongly rejected in favor of a hybrid specification containing lagged as well as future inflation, a result that has not been adequately explained in a consistent theoretical fashion see Roberts [1997], Galı and Gertler [1999], Rudd and Whelan [2005], Fuhrer [2006], Whelan [2007] and Rumler [2007] amongst a voluminous literature. Worse still when DC is relaxed by allowing persistent distortionary shocks,[5] optimal policy amounts to a form of price-level targeting Woodford [2010]. Therefore, inflation inherits *negative persistence*.

This contrasts with best practice among inflation-targeting Central Banks who practice so called *'coarse-tuning'* Lindbeck [1992]. They realize that inflation possesses intrinsic persistence so that they cannot hit inflation and output targets in every period. Instead they practice so called *inflation forecast-targeting* Kohn, Svensson [2010] and Svensson [2012]. This is where policy and projections for future policy are adjusted to yield a desirable expected path for inflation and real activity, consistent with *medium term* stability. This is usually defined as forecast inflation and output gap sufficiently close to target after a time frame of 18 months to 3 years[6].

A related challenge confronting New Keynesian modeling is the policy persistence puzzle. Interest rates are highly persistent - much more so than underlying shock processes. This means that estimated Taylor rules require coefficients on lagged rates near unity to purge serial correlation and represent an optimizing relationship see Coibion and Gorodnichenko [2012] and Vázquez et al. [2013][7] Attempts to explain

[4]It is worth noting that many studies are able to reject the null of no persistence in inflation even when they have sufficient power to uncover statistically significant changes in persistence across policy regimes. Although, there is considerable heterogeneity in inflation persistence across sectors macroeconomic persistence is not a figment of aggregation bias.

[5]Distortionary shocks effect the wedge between actual and efficient output. They enter into the NKPC where they are often called 'cost-push' shocks.

[6]Studies with VARs and policymakers wisdom suggest that it takes between 18 months and two years for a change in monetary policy to have its maximum impact on inflation. This result seems to be robust across changes in policy regimes; Orlowski [2000] (see p 315-320), Batini and Nelson [2001] and Gerlach and Svensson [2003]. On its website the Bank of England advises the general public that: "Monetary policy operates with a time lag of about two years." http://www.bankofengland.co.uk/monetarypolicy/Pages/overview.aspx However the Bank publishes forecasts three years ahead and frequently talks about "inflation returning to target by the three year horizon" consistent with a longer view of the stabilization and empirical work by Havranek and Rusnak [2013] period.http://www.bankofengland.co.uk/publications/Pages/inflationreport/infrep.aspx Practices are similar at other leading inflation targeting Central Banks.

[7]These papers control respectively for non-rational expectations on the part of the policymaker and data revisions. See also Rudebusch [2002], Petra [2004], Rudebusch [2006], Carrillo et al. [2007] and Conraria et al. [2014]. The acclaimed Norges Bank (the Central Bank of Norway) chooses to insert a substantial interest rate stabilization term- ad hoc with respect to its policy mandate- into its loss function used to derive optimal policy- in order to generate policy rate predictions consistent with credible application of inflation forecast targeting see Bergo [2007] and Holmsen et al. [2007]. Other Central Banks swerve around this problem by using subjective judgments or information from

why this might be optimal have so far proven unconvincing.

The three equation framework was designed explicitly to address issues with optimal monetary policy and its effects upon inflation and real activity. New Keynesian theory is failing the test of policy relevance. Chari et al. [2009] were right: New Keynesian models are not yet fit for purpose. No one could resolve these problems in a single thesis. My contribution here is twofold. First of all, I explain how two of these unfortunate results; lack of persistence and divine coincidence arise from a faulty linear approximation at a point unrepresentative of the dynamics elsewhere. Secondly, I formalize a general equilibrium concept that can be used to understand the probabilistic behavior of an entire DSGE model. I then show how this can be used to derive approximations to the dynamic behavior of the economy close to its steady state.

This research is related to two strands of recent macroeconomic literature. In monetary economics it has been known since at least Ascari and Rankin [2002] that the dynamics of the Calvo model in particular are very different when linearized around a positive rather than zero trend inflation. In particular there is endogenous inflation persistence stemming from the Phillips curve- Cogley and Sbordone [2008] and Cogley et al. [2010]. Alves [2014] shows how a non-zero trend inflation causes a welfare loss and therefore contradicts Divine Coincidence. Trend inflation appears to make the economy more volatile and more difficult to stabilize see Ascari and Ropele [2009], Coibion and Gorodnichenko [2011] and Ascari et al. [2015]. Ascari and Sbordone [2014] surveys this burgeoning literature.

The second strand is the risky steady state literature. This has come out of macro-finance and the problem of portfolio selection, where it is clear the non-stochastic solution, in which assets have the same return, is unsatisfactory for analyzing changes in portfolio allocation and returns. Well-known formulations include Juillard and Kamenik [2005], Coeurdacier et al. [2011], Michel [2011] and de Groot [2013]. None of them are mathematically precise in terms of definition but have proven useful in application.

My contribution is to show that we can generate trend inflation behavior by adding

futures rates or the yield curve which have expectations of policy persistence built-in Ang et al. [2007] and Hamilton et al. [2011].

risk into our understanding of the steady state, even if there is no trend in inflation. Secondly, I formalize concepts in the risky steady state literature that might help to develop more accurate techniques to understand stochastic equilibrium in financial markets. I also show that the problem with non-stochastic steady state approximations relates to bifurcations points unrepresentative of local dynamics rather than risk per se. Finally, I hope that my work here will act as a bedrock for future progress on the many puzzles with monetary policy and the business cycle.

Chapter 2

New Keynesian Framework

This section focuses on the Calvo pricing framework the most popular staggered pricing model. I start by analyzing the general non-linear model without specifying the trend inflation rate coming from the Central Banks inflation target. I investigated several avenues to exposit the mechanisms of the model or because they will prove profitable later on. I then adopt the popular approach of log-linearizing around its zero inflation non-stochastic steady state. I demonstrate problems with persistence and identification. For comparison, I then introduce two models one the Keynesian Rotemberg pricing model and the New Classical Lucas imperfect information model.

2.1 Household's Problem

This section is divided into two blocks, the first subsection concerns how the household allocates its resources between aggregate consumption and leisure - with consequences for aggregate supply and demand. The second considers how the household determines its consumption bundle. It helps to make important aggregation concepts easier to understand.

2.1.1 Aggregate Allocation

There is a representative household which solves the following problem:

$$\max_{C_t, l_t} \mathbb{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \Big[u(C_T) - \varphi_T \nu(l_T) \Big] \psi_T \tag{2.1}$$

subject to the Budget Constraint:

$$P_tC_t + B_{t+1} = (1 + i_{t-1})B_t + P_tW_tl_t + T_t + \int_i \Pi_t(i)\, di \tag{2.2}$$

β is the discount factor, C refers to aggregate consumption whilst l is labor supply. B refers to the holding of one period risk free nominal bonds. i_t is the risk-free nominal interest rate paid at the end of period t on the bond. P is the price level - bonds are the numeraire here. W is the real wage. T is a lump sum transfer between the government and the household. Finally $\Pi(i)$ is profit from an individual firm i given by

$$\Pi_t(i) = (1 + \tau_t(i))\, p_t(i)\, y_t(i) - W_tl_t(i) \tag{2.3}$$

here $\tau_t(i)$ is a production subsidy that may vary across firms' and time. For simplicity let the government budget always balance so

$$\int_i \tau_i\, p_t(i)y_t(i)\, \mathrm{d}i = T_t \tag{2.4}$$

Fiscal policy is not the focus of this paper. The two instruments will only be to undertake thought experiments related to welfare and optimal policy. Unless otherwise stated they will be turned off so[1]

$$\tau_t(i) = \tau_t = T_t = 0 \tag{2.5}$$

Note in a stochastic environment firms need not make the same profits even with a symmetric equilibrium. Since when price rigidity is introduced firms with the same demand curve will charge different prices depending on when they last reoptimized and can therefore make different levels of profit. In fact without arbitrary restrictions on shock size there is no way of insuring non-resetters have positive profit expectations and therefore a non-negative stock price. To swerve around this problem imagine

[1]This economy is Ricardian in the sense that the issue of government debt would not influence other variables in the economy- in particular the intertemporal pattern of consumption provided the subsequent equilibrium conditions all continued to hold and the following long-run budget constraint always held

$$\mathbb{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \left[\int_i \tau_i\, p_t(i)y_t(i)\, \mathrm{d}i - T_t \right] \psi_T = 0$$

there are a continuum of households who each own one firm and an efficient long-term insurance market with contracts set arbitrarily far back in the past (so that there is no conditioning on pricing history.) This is effectively an optimal timeless insurance contract in analogy to optimal timeless monetary policy Woodford [1999]. In this case the market will absorb all idiosyncratic risk associated with non-optimal price setting behavior.[2] Weber [2014], Gorodnichenko and Weber [2016], D'Acunto et al. [2017] and Ozdagli and Weber [2017] provide supportive empirical evidence of the effect of nominal rigidity on firm profitability and other aspects of corporate behavior, as well as more refined theory.

ψ_T is a positive demand shock. The budget constraint states that the uses for nominal income- consumption and saving- must be equal to the sources of income- wealth, labor and dividend income.

Consumption is desirable but working is undesirable so u and ν are increasing. u is concave to incentivize consumption smoothing whilst ν is convex to encourage workers to take leisure. An Inada condition on u, a restriction on any shock processes, a zero wealth condition and a transversality condition support a well-behaved solution They are as follows:

$$\lim_{C \to 0^+} u_c(C) = \infty \tag{2.6}$$

$$B_t = 0 \tag{2.7}$$

$$\sup_{\{C_t, l_t\}} \mathbb{E}_t \sum_{T=t}^{\infty} \beta^{T-t} |u(C_T) - \varphi_T \nu(l_T)| \, \psi_T < \infty \tag{2.8}$$

$$\lim_{T \to \infty} \mathbb{E}_t \frac{B_T}{P_T} u_c(C_T) \geq 0 \tag{2.9}$$

Equation (6) is a "no-starvation" condition, it ensures the agent will always choose to consume even though working is costly. Equation (7) stops the representative agent living off their savings. It is the bond market clearing condition of the representative agent economy when there are no taxes or transfers- as specified by equation (5).[3]

[2] This 'cut-round' is necessary for all existing results from models that use unbounded shocks and non-optimal behavior from the planners point of view including sticky information and Taylor pricing. Alternatively I could have the social planner use lump sum taxes to correct portfolio effects of nominal rigidity. The case of state dependent pricing will be discussed later in Chapter 2.6 in the specific case of Rotemberg pricing.

[3] This is not a substantive restriction. If we had non-zero lump sum taxes then the economy would be Ricardian so the level of government debt would be irrelevant since government bonds would not affect net worth as famously demonstrated in Barro [1974].

Together equations (6) and (7) ensure strictly positive labor supply. Equation (8) is required for a well-defined solution to the consumers problem amenable to recursive analysis. Finally, equation (9) is a "No-Ponzi" condition; it forces the agent to honor the present value of their debts. Due to monotonicity of the utility function u the constraint will always bind with equality and if it were left out the agent would demand to borrow an infinite amount and never repay. This condition will feature in the proof of existence of stochastic equilibrium in chapter 5. Therefore it is worth noting that Giglio et al. [2015], Giglio et al. [2016] and Fesselmeyer et al. [2016] supply model independent evidence in favor of transversality conditions.

The first order conditions for the households are as follows:

$$u_c(C_t) = (1 + i_t)\beta \mathbb{E}_t U_c(C_{t+1}) \frac{\psi_{t+1}}{\psi_t} \frac{P_t}{P_{t+1}} \tag{2.10}$$

$$u_c(C_t)W_t = \varphi_t \nu_l(l_t) \tag{2.11}$$

Equation (7) the so-called Euler equation specifies the path for optimal consumption, whilst equation (8) which equates the marginal costs and benefits of working, yields the labor supply curve.

Dynamic properties will be characterized in terms of two parameters

$$\sigma = \frac{-Cu_{cc}}{u_c} > 0$$

$$\eta = \frac{l\nu_{ll}}{\nu_l} > 0$$

The parameters measure respectively the concavity of consumption utility and the convexity of the disutility from work. σ is the inverse of the inter-temporal elasticity of substitution - the consumer's willingness to shift consumption across time periods, Similarly η is the Frisch elasticity of labor supply and is inversely related to the propensity for inter-temporal labor substitution.[4] Where beneficial I will use specific

[4] σ is also the coefficient of relative risk aversion but this interpretation is not relevant initially as the solution concept used in this section taking first order perturbations about a non-stochastic steady state - yields certainty equivalence. Later on when I consider stochastic steady states - the distinction between risk aversion and intertemporal substitution could become operative. However, I leave this avenue to future inquiry.

functional forms

$$u(C) = \begin{cases} \frac{C^{1-\sigma}}{1-\sigma}, & \sigma \neq 1 \\ \\ \log(C), & \sigma = 1 \end{cases} \tag{2.12}$$

$$\nu(l) = \frac{l^{1+\eta}}{1+\eta} \tag{2.13}$$

Finally, for simplicity I start with no capital hence the budget and resource constraints are equivalent to

$$C_t = Y_t \tag{2.14}$$

which will naturally pass down to the output of individual goods and firms as $c_t(i) = y_t(i)$. This completes the characterization of consumer preferences over aggregates goods.

2.1.2 Individual Product Demand

The purpose of this part is to derive the demand for an individual variety. For convenience let the measure of the firms be a unit continuum. The firms are monopolistically competitive. This ensures that firms face a meaningful pricing decision, avoiding the case of unbounded sales possible under perfect or simple Bertrand competition. Initially there will be no idiosyncratic shocks. Later on when I relax this assumption more exposition will be provided. Under monopolistic competition firms produce differentiated products that are imperfect substitutes and consumers prefer variety. This implies that aggregate consumption has to be a non-linear function of the underlying individual varieties. This contrasts with labor where if we broke up the representative household into constituents j then the aggregation would be linear with $l = \sum_{j=1}^{J} l_j$

The consumption objective (utility from consumption) is

$$U = \left[\int_0^1 c_t(i)^{\frac{\theta-1}{\theta}} \, di \right]^{\frac{\theta}{\theta-1}} \tag{2.15}$$

where $\theta > 1$ is required for a well-behaved problem. The household's problem is to maximize consumption utility subject to an expenditure constraint.

$$P_t C_t = \int p_t(i)\, c_t(i)\, di \tag{2.16}$$

For any two varieties i and i' this yields relative demand.

$$\frac{c_t(i)}{c_t(i')} = \left(\frac{p_t(i)}{p_t(i')}\right)^{-\theta} \tag{2.17}$$

Now a little manipulation and then integration with respect to i' yields:

$$\int p_t(i')\, c_t(i')\, di' = \int c_t(i)\, p_t(i)^{\theta} p_t^{1-\theta}(i')\, di' \tag{2.18}$$

Guessing and verifying yields the demand system price level pair

$$P_t = \left[\int_0^1 p_t(i)^{1-\theta} di\right]^{\frac{1}{1-\theta}} \tag{2.19}$$

$$c_t(i) = \left(\frac{p_t(i)}{P_t}\right)^{-\theta} C_t \tag{2.20}$$

In the benchmark no capital world this becomes

$$y_t(i) = \left(\frac{p_t(i)}{P_t}\right)^{-\theta} Y_t \tag{2.21}$$

2.2 Firms' Problem

Throughout the paper this will be the source of monetary non-neutrality and the font of the Phillips curve. A novel feature Generalized Stochastic Mean is introduced which gives insight into the cause of monetary non-neutrality and the significance of the zero inflation non-stochastic steady state. The rest is standard, firms maximize profits in the choice of factors and also in prices when they are given the chance to reoptimize. In order to profit maximize they must cost minimize- for the simple proof consult Varian [1992].

2.2.1 Cost Minimization

For simplicity there is only one factor of production: labor purchased on a competitive market. The production function takes the form

$$y_t(i) = f(l_t(i)) \tag{2.22}$$

where f_t must always be weakly concave ($f_{ll} \leq 0$) to ensure a well-behaved solution and the time script permits stochastic developments in technical efficiency. The problem is as follows:

$$\min_{l_t(i)} W_t l_t(i) \tag{2.23}$$

subject to the production constraint

$$f(l_t(i)) = \bar{y}_t(i) \tag{2.24}$$

The Lagrange multiplier gives the real marginal cost of production paid by the firm. Hence we can solve for real marginal cost

$$MC_t(i) = \frac{W_t}{f'(l_t(i))} \tag{2.25}$$

Note that in general marginal costs will differ across firms as they depend on the marginal product of labor which is firm specific because of variable returns to scale. In the main text I will work with a linear production technology which ensures all firms will have the same marginal costs simplifying analysis considerably

$$MC_t(i) = MC_t = \frac{W_t}{A_t} \tag{2.26}$$

in particular the real marginal cost is the ratio of the real wage W_t to aggregate technical efficiency term A_t, from the RBC model. In the benchmark New Keynesian model firm level productivity shocks either do not exist or have been averaged out by a law of large numbers.

2.2.2 Optimal Price Setting

Calvo pricing is the most popular approach to inject nominal rigidity into a DSGE model. Reoptimization is governed by a stochastic process common across firms. With probability $1-\alpha$ each firm is free to reset its price (at no cost), whilst with probability α it is not allowed to reoptimize and meets demand at its existing price. The firms' resetting prices maximize the expected present value of profits through the lifetime of the price as follows:

$$\max_{p_t^*(i)} \mathbb{E}_t \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \left[\frac{p_t(i)}{P_T} y_T(i) - C(y_T)(i) \right] \tag{2.27}$$

subject to demand and market clearing constraints faced by all firms:

$$y_t(i) = \left(\frac{p_t(i)}{P_t} \right)^{-\theta} C_t \tag{2.28}$$

Here

$$Q_{t,t+k} = \beta^k \frac{U_c(C_{t+k})}{U_c(C_{t+k})} \Pi_{t,t+k}^{\theta}$$

represents the real stochastic discount factor (SDF), it is the risk-adjusted present value of future consumption k periods ahead which depends on the gross rate of inflation

$$\Pi_{t,t+k} = \frac{P_{t+k}}{P_t} = (1+\pi_{t+1}) \cdots (1+\pi_{t+k})$$

is between today time t and a future time $T > t$.

Unlike other infinite horizon problems encountered in economics, in general, the reset price problem is *non-recursive* as the firms' choice variable p_t^* depends not just on the state of the economy next period $t+1$ but its whole future[5]. This is the source of

[5] This detail is specific to Calvo pricing in Taylor economies the optimization problem of each firm is finite horizon and in state dependent models firms' optimize each period which gives rise to a recursive structure if there are state variables present and otherwise collapses into a static problem. Later on, the model will be reformulated to create a recursive structure using variables that depend on all past and expected future states of the economy. Recall that a recursive problem is one where the optimization can be rendered as a dynamic programming problem

$$\max_{z_t} V_t(z_t, \psi_t) = u(x_t, z_t, \psi_t) + \beta \mathbb{E}_t V_{t+1}(z_{t+1}, \psi_t)$$

where V_t is today's objective function which depends upon today's value of z_t the state variable (one whose expectation is set in the previous period so $z_{t+1} = E_t z_{t+1}$), a stochastic term ψ_t governed by a rule similar to (5) to ensure a well-defined solution and the jump or control variable x_t which is determined in each period. u and $\mathbb{E}_t V_{t+1}$ are known respectively as the instantaneous and continuation pay-off. In the canonical consumption-savings problem which underpins the RBC framework z

endogenous persistence in the benchmark New Keynesian framework. Non-recursive optimization problems are common to all settings in which nominal rigidity arises through some firms not changing their price every period[6]. The first order condition is

$$\mathbb{E}_t \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \frac{\partial y_T(i)}{\partial p_t(i)} \Big[MR_T(y_T(i)) - MC_T(y_T(i)) \Big] = 0 \tag{2.29}$$

It states that optimal pricing sets a weighted stream of marginal revenues equal to a weighted stream of marginal costs, which in turn implies a similar relationship between (real) price and marginal costs.

$$\mathbb{E}_t \sum_{T=t}^{\infty} (1-\alpha)^{T-t} Q_{t,T} \left(\frac{p_t^*}{P_T} \right)^{-\theta} Y_T \left[\frac{p_t^*}{P_T} - \frac{\theta}{\theta - 1} MC_T(y_T(i)) \right] = 0 \tag{2.30}$$

By strict concavity of the optimization problem the equilibrium price $p_t^*(i)$ is unique for each firm so there is a unique optimal reset price each period p_t^*. Therefore by a law of large numbers argument, see Anderson et al. [1991], the price level evolves as follows.

$$P_t^{1-\theta} = \alpha P_{t-1}^{1-\theta} + (1-\alpha)(p_t^*)^{1-\theta} \tag{2.31}$$

The persistence of the price level depends on α the degree of rigidity.

2.2.3 Flexible Price Equilibrium and Zero Inflation Non-Stochastic Steady State

This subsection clarifies the coincidence between the ZINSS and the flexible price equilibrium - a result crucial to the whole paper. Mathematical content including extensive derivations can be found in Appendix 7.1. Under monopolistic competition we find that the optimal price is a mark up $m = \theta/(\theta - 1)$ over marginal costs so in real terms

$$\frac{p_t^f}{P_t} = \frac{\theta}{\theta - 1} MC_t \left(\frac{p_t^f}{P_t} \right) \tag{2.32}$$

Particular, to the flexible price equilibrium is that whatever has happened in the past all firms have the same price $p_t^f = P_t$. Also the real marginal cost MC and markup m are common across firms with $MC_t = \bar{MC} = (\theta - 1)/\theta = 1/m$. The connection

is capital, x consumption and u the households contemporaneous consumption, which would depend only on consumption.

[6] It would also carry over to common indexation schemes.

between the optimal flexible price p_t^f and the optimal reset price p_t^* underpins the equivalence between the zero inflation non-stochastic steady state and the flexible price equilibrium. It takes the form of a *Generalized Stochastic Mean* (GSM) defined as follows

Definition 1. A GSM $\mathfrak{M}$ over a collection X ordered by $i \in \mathbb{N}$ whose elements take on values $x \in \mathcal{X} \subseteq \mathbb{R}^n$ is a function stochastic process pair $(\mathcal{M}(X), \mathcal{S}(X))$ such that

1. $\mathcal{M}(\bar{x}, \bar{x}, \cdots, \bar{x}) = \bar{x},$ for any sure (non-stochastic) sequences $\langle \bar{x}, \bar{x}, \cdots \rangle$

2. $\mathcal{M}$ is strictly increasing, measurable and continuous on $\mathbb{R}^\infty$ the space of convergent sequences and is weakly measurable with respect to the stochastic process $\mu(\langle X \rangle)$.[7]

Lemma 1. *The optimal reset price p_t^* is a GSM of the optimal flexible price p_T^f for the duration of the contract $T \geq t$.*

Corollary 1. *A non-stochastic steady state is equivalent to a flexible price equilibrium*

Proof. From condition [ii] we have $p_t^* = \mathcal{M}(\langle p_{NSS}^f, p_{NSS}^f, \cdots \rangle) = p_{NSS}^f, \ \forall T$ □

Therefore at the non-stochastic steady state the reset price constraint is not binding. Furthermore, any variable that measures the distance between the classical flexible price world and its sticky counterpart will be minimized at the non-stochastic steady state. This point will prove crucial when I come to approximation of the model.

2.2.4 Future Marginal Costs and Nominal Rigidity

The averaging result carries over to real marginal costs.

Corollary 2. *The optimal real value of the current reset price p_t^*/P_t is a GSM of the real marginal cost of the flexible price firm $MC_T^*(i)$ for the duration of the contract $T \geq t$.*

[7]Results in this subsection apply also to the case of Taylor contracts where the topological vector space itself would collapse into a Euclidean manifold - reflecting the finite contract horizon. Results would extend over to any reflexive space- informally this is a space where the set of possible ensemble averages and the set of possible realizations coincide see Conway [2013]

Proof. $\mathbb{E}_t \partial p_t^* / \partial MC_\tau^* = \mathbb{E}_t [P_\tau / P_t \cdot \partial p_t^* / \partial p_\tau^f] > 0$ $\square$

The gross rate of inflation $\Pi_{t,T} = P_T/P_t = (1+\pi_{t+1}) \cdots (1+\pi_T)$ is between today time t and a future time $T > t$. The following remark locates the proximate source of *policy effectiveness* in terms of the failure of the classical dictum of independence between real and nominal prices

Remark 1. *Monetary* policy effectiveness *applies because the real reset price depends on expected future inflation.*

It also overturns the neoclassical principal that only unexpected inflation affects real quantities which are determined by real prices. Enlisting the functional form assumption leads to a familiar expression for the optimal reset price equal to a mark-up over the weighted average of expected marginal costs over the infinite future denoted by WMC_t^*

$$\frac{p_t^*}{P_t} = m\mathbb{E}_t \sum_{T=t}^{\infty} w_T MC_T^* = mWMC_t^* \tag{2.33}$$

where the weight w_T can be decomposed into a real component w^r that reflects the scale of real output in future periods relative to today and nominal component $w^n = 1/\Pi_{t,T}$ reflecting the effect of inflation through the formulation $w_T^r / (\sum_{T=t}^{\infty} w_T^r w^n)$ where $w_T^r = \alpha^{T-t} Q_{t,T} g_{t,T}^y$ where the first two terms reflect stochastic factors already discussed and $g_{t,T}^y = \Pi_{t,T}^{\theta} Y_T / Y_t$ is the expected growth rate of sales up to time T conditional on not having reset its price. Combination with the price level evolution (2.31) yields a formulation for the non-linear marginal cost Phillips curve

$$\pi_t = \frac{(mWMC_t^*)^{1/(\theta-1)}}{1 + \alpha(mWMC_t^*)^{\theta-1}} \tag{2.34}$$

It is more convenient for the purpose of derivation however to separate the right hand side into numerator and denominator $p_t^*/P_t = m\aleph_t / \beth_t$ where

$$\aleph_t = \mathbb{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \Pi_{t,T}^{\theta} u_c(C_T) Y_T MC_T \tag{2.35}$$

$$\beth_t = \mathbb{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \Pi_{t,T}^{\theta-1} u_c(C_T) Y_T \tag{2.36}$$

This is because the numerator and denominator both possess recursive structures shown below.

$$\aleph_t = u_c(C_t)Y_tMC_t + \alpha\beta\mathbb{E}_t(1+\pi_{t+1})^{\theta}\aleph_{t+1} \tag{2.37}$$

$$\beth_t = u_c(C_t)Y_t + \alpha\beta\mathbb{E}_t(1+\pi_{t+1})^{\theta-1}\beth_{t+1} \tag{2.38}$$

2.3 Price Dispersion and Aggregation

To understand the dynamics of the system it is common to log-linearize, for business cycle interpretation we need to choose a point around which to carry out this perturbation that could be interpreted as a long-run equilibrium. Economists have always used the *non-stochastic steady state* that would prevail if the economy were never subject to shocks or expected to be so. The non-stochastic steady state is equivalent to the flexible price equilibrium. The point of this paper is that this approach is misguided providing an erroneous interpretation of business cycle dynamics.

The New Keynesian framework differs from Neoclassical models by preventing every firm re-optimizing prices in every period. This allows for the possibility of *price rigidity* where today's price level contains reset prices from previous periods, as well as the current optimal reset prices. This means there can be *price dispersion* with implications for resource allocation and welfare.

We characterize price dispersion using the demand aggregator

$$\Delta = \int_i \left(\frac{p_i}{P}\right)^{-\theta} \mathrm{d}\mu_i \tag{2.39}$$

It appears in the market-clearing condition

$$\Delta_t C_t = A_t L_t \tag{2.40}$$

Under Calvo pricing Δ evolves according to the following relationship:

$$\Delta_t = (1-\alpha)\left(\frac{p_t^*}{P_t}\right)^{-\theta} + \ \alpha(1+\pi)^{\theta}\Delta_{t-1} \tag{2.41}$$

Using equation (2.31) to eliminate the reset price we find:

$$\Delta_t = \frac{[1-\alpha(1+\pi)^{\theta-1}]^{\frac{\theta}{\theta-1}}}{(1-\alpha)^{\frac{1}{\theta-1}}} + \alpha(1+\pi)^{\theta}\Delta_{t-1} \tag{2.42}$$

Price dispersion is a persistent process with the degree of persistence increasing in the degree of price rigidity α. In Chapter 4 I will discuss the restrictions the consumers' problem imposes on inflation. The following property of the log-linear form is remarkable.

Lemma 2. *Around the non-stochastic steady state* $(\pi, \Delta) = (0, 1)$ *the log-linear approximation* $\hat{\Delta}_t = 0$ *for all values of inflation* π_t.

Proof. Log-linearizing the system reveals that

$$\hat{\Delta}_t = \frac{\alpha\theta(1+\pi)^{\theta-2}}{\Delta(1-\alpha)^{\frac{1}{\theta-1}}}[(1+\pi)\Delta(1-\alpha)^{\frac{1}{\theta-1}} - (1-\alpha(1+\pi)^{\theta-1})^{\frac{1}{\theta-1}}]\hat{\pi}_t + \alpha(1+\pi)^{\theta}\hat{\Delta}_{t-1}$$

substitution of the steady state values $(\pi, \Delta) = (0, 1)$, along with $\hat{\Delta}_{t-1} = 0$ which defines a non-stochastic steady state and $\hat{\pi}_t \equiv \pi_t - \pi = \pi_t$ completes the proof. □

2.4 Policy Rule

This section explains the monetary policy rule used throughout this thesis. It also contains a more in depth discussion of the problems encountered with modeling monetary policy in a DSGE framework. Novel points include motivations for the Taylor rule and non-linear rules. In practice none of the models used in empirical simulations can be viewed as optimizing a suitable loss function.The most well-known is the so-called Taylor rule. The alternative is to specify a money supply rule. The most popular money supply rule was the McCallum rule, see McCallum [1988]. The McCallum rule is essentially a flexible version of the monetary targeting rules advocated by monetarists like Milton Friedman, and applied by Western policymakers as part of disinflationary efforts during the late 1970s and early 1980s. In Britain and the United States monetary targeting was eventually abandoned owing to perceived instability in the money demand function which impaired implementation.

It now seems that this appearance was a figment of a mismeasuring the opportunity cost associated with the new money substitutes on offer as a consequence of the financial deregulation occurring around that time- see Ericsson et al. [1998], Ireland [2009], Barnett [2012], Ball [2012], Lucas and Nicolini [2015]. In the stylized models shown here money demand is inconsequential to policy because of the presence of frictionless financial markets as shown by Woodford [1998]. The empirical application of the result is doubtful however, as it appears empirical specifications including money perform better than those with just interest rates, see Belongia and Ireland [2014]. This seems to be particularly important in periods of very low interest rates - where the interest rate channel breaks down but policy still appears to be effective, see among others Ueda [2012], Kapetanios et al. [2012], D'Amico et al. [2012], Swanson and Williams [2014] and Gilchrist et al. [2015].

The problem appears to be the restriction on the class of financial market inefficiency imposed in benchmark models, as explained in Jarrow [2013]. It has been suggested that these distortions may have substantial effect on monetary policy propagation in normal times, consult Jiménez et al. [2014], Gertler and Karadi [2015], Nelson et al. [2015], and for optimal policy Chadha et al. [2014], Ellison and Tischbirek [2014] and de Groot [2014]. Nevertheless, results tend to be model specific and may also depend on other aspects of the policy and regulatory environment, see Svensson [2014]. Furthermore, macro-prudential concerns do not seem to have played a major role in monetary policy determination, see Fuhrer and Tootell [2008] and Rotemberg [2014] for evidence and discussion of institutional arrangements. There are several example of leading policymakers arguing Central Banks should not try to burst bubbles for example Bernanke and Gertler [2001] and Posen [2006]. These concerns support my decision to work with a benchmark efficient financial markets setup.[8] The following is a popular modern formulation, I discuss the relation to the original proposal of Taylor [1993a] in Appendix A:

$$i_t = \bar{\imath}_t + a_\pi \tilde{\pi}_t + a_y \tilde{y}_t \tag{2.43}$$

[8] It is worth noting that even though they are associated with inflation targeting Taylor rules seems to explain policy-making just as well even when the Central Bank professes to be following a different policy regime such as money targeting, see Bernanke and Mihov [1997] and Clarida et al. [1998]. This chimes with the recollections of Goodhart [1988] that even when publicly espousing monetary targets policy advice would be given in terms of the nominal interest rate rather than the money supply. These points increase my confidence in applying an inflation targeting model to the United States and other nations where there is no explicit inflation targetter.

where $a_\pi > 1$ and $a_y \geq 0$. $\tilde{\pi}$ refers to the inflation gap the difference between observed inflation π_t and its target $\bar{\pi}$. $\tilde{y}_t$ is the output gap defined as the difference between the actual output y_t and its natural or potential rate $\bar{y}_t$. This potential rate is calculated by either deterministic time trend or a stochastic trend derived from a more sophisticated filter such as Beveridge and Nelson [1981], Hodrick and Prescott [1997] or Baxter et al. [1999], there are now standard settings, see Ravn and Uhlig [2002] and Christiano and Fitzgerald [2003][9]. In the context of the benchmark model and its variants, this potential rate is associated to the flexible price equilibrium derived previously. $\imath_t$ is the natural nominal rate of interest. It comprises a natural real rate $\bar{r}_t$ and the inflation target $\bar{\pi}_t$. Although, not present in the original paper it is common to allow these two long run parameters to time vary. There are two main motivations, firstly if the model is estimated, all but the simplest filtering processes create a time varying estimate for the inflation target.

The coefficient restrictions form the essence of Taylor's contribution. $a_\pi > 1$ requires that the nominal interest rate should rise more than one-for-one (so the real interest rate increases) with inflation to help ensure a stable solution[10]. $a_y \geq 0$ constitutes an allowance but not a requirement for interest rates to smooth the output gap - consistent with the 'dual mandate' to smooth fluctuations in the real economy as well as inflation common to all major inflation targeting nations - consult Svensson [2010], Svensson [2012] and Rotemberg [2014] for more detailed discussion of the United States experience. The crucial difference between Taylor's original normative rule and its implementation in subsequent DSGE models lies in the policy horizon.

Taylor envisaged a forward-looking rule designed to gauge whether or not the current level of interest rates was consistent with stabilizing inflation at its long run average level over subsequent periods given the current state of the economy summarized by $\tilde{y}_t$. Modern DSGE follow this convention also. Forward-looking intent can be seen by how inflation was incorporated into the original rule. Taylor used the annual inflation

[9]It is necessary to filter the raw data to remove the effect of the trend in technical progress as it is not the focus of interest in this stylized model I will continue to ignore it throughout the paper.

[10]This condition is necessary and sufficient for a stable solution in the pure inflation targeting case (uncharitably referred to as 'inflation nuttery' if you recall) corresponding to $a_y = 0$. Otherwise (when $a_y > 0$) the condition can be weakened since the positive relationship between output and inflation implied by the New Keynesian Phillips curve allows the combined effect of the inflation and output gaps to drive up the real rate even for some $a_\pi < 1$. However, the intuition that the real rate must be expected in response to an increase in inflation remains valid. Consult Woodford [2011a] for this result which is derived in the context of a contemporaneous response to shocks.

rate $\pi_t = (P_t - P_{t-4})/P_{t-4}$ as opposed to the conventional contemporaneous output gap measure $\tilde{y}_t = (y_t - \bar{y}_t)/\bar{y}_t$ where y_t was quarterly output and $\bar{y}_t$ was determined from a simple linear trend over the whole data period.[11] He argued that "the lagged inflation rate [was] serving as a proxy for expected inflation." This logic was perfectly correct with respect to the staggered wage setting model (Taylor [1979]) he was using to study policy implications of monetary rules in Taylor [1993b] which grounded his normative analysis. I demonstrate this point in appendix A. However, I will show it is not in fact the case with the benchmark log-linearized New Keynesian model with Calvo pricing.

There are no lags or leads in the relationship so contemporaneous economic developments determine current policy. This is intuitive, there is no benefit to conditioning on past variables in a forward-looking model[12]. In the particular but instructive case with no persistence to exogenous shock processes the future of a forward-looking system will be identical in expectation to the steady state. In a forward-looking system one instrument per period should be sufficient to implement optimal policy.

Note, we are not required to believe that interest rates are determined simultaneously with output and inflation. In subsequent sections interest rates will be determined before output and inflation are realized. As the rule is linear with rational expectations the expectation errors will pass into the white noise error term, providing in fact the main rationale for the regression error itself. I interpret the policy rule as a description of actual policy for a Central Bank with rational expectations[13]. Hence, I view the presence of serial correlation and a significant lagged interest rate term as a rejection of Taylor rule as a description as a model of real-world policy-setting.

Finally, the Taylor rule or any linear policy rule is best viewed as a local approximation to an unknown global regime. There is plausible evidence for non-linear rules and

[11] Taylor only considered the Great Moderation period 1984:1- 1992:3. His trend growth rate was 2.2 percent on an annualized basis and he used 2 percent for the real interest rate.

[12] The system is 'forward-looking' in the sense that agents have rational expectations and there are no state variables.

[13] There is some controversy about whether Central Banks do display rational expectations see Romer and Romer [2008] and Ellison and Sargent [2012]. However, the conclusion of excess persistence appears robust to and even strengthened by the use of internal forecasts understood to be the best forecast of the evolution of the economy see Coibion and Gorodnichenko [2012]. In any case the significance and substantiveness of persistence appears too large to be explained solely by deviations from rational expectations.

some theoretical rationale- see for example Blinder et al. [2000], Meyer et al. [2001], Ruge-Murcia [2003],Surico et al. [2007], Chevapatrakul et al. [2009] and Neuenkirch and Tillmann [2014]. Here the non-linear Taylor rule could serve the role of ruling out the boundary solution case where the representative agent seeks no leisure. This is because with the popular functional form in place there is no equivalent of the Inada condition to prevent the individual taking no leisure. A functional form like Cobb-Douglas with arguments leisure and consumption would provide such a condition. However, I have not done so as this is not common in the macroeconomic literature. Leaving this technicality aside. It is now possible to specify the Calvo framework in full.

2.5 Formalization Calvo Model

This subsection gives the canonical state space form of the non-linear Calvo model. It also spells out the mathematical assumptions that will be in place throughout the paper to support the formal analysis.

Assumption 1. *The Calvo model described by equations (2.1)-(2.43) and the two alternative structural models subsequently in Section 2.7 presented throughout this paper possess a canonical form* $\mathbb{E}_t Z_{t+1} = f(Z_t, \gamma, U_t^{\mathcal{S}})$ *with the following properties*

i Z *and* $U^{\mathcal{S}}$ *are respectively* $k \times 1$ *and* $k' \times 1$ *dimensional vectors that can be decomposed into jump and state variables* $Z = (Z^j_{k_0\times 1}, Z^s_{k_1\times 1})$, $U = (U^j_{k'_0\times 1}, U^s_{k'_1\times 1})$ *such that* $k_0 \leq k'_0$ *and* $k'_1 = k'_0$

ii The interior of $\mathcal{Z}$ *and* $U^{\mathcal{S}}$ *forms smooth manifolds*

iii Z^j *is stochastically non-singular so that for any* $1 \leq i \leq k_0$ *there is no non-stochastic function* $\tilde{f}$ *such that* $Z_{i,t} = \tilde{f}(Z_{-i,t})$. U^j *will be mutually independent*

iv $\mathbb{P}(Z^j_{t+1} \in B|(Z_t, U_t^{\mathcal{S}})) > 0$ *for all* $B \in \Sigma^j$ *where* Σ^j *is the* σ*-algebra of the vector of jump variables* Z^j

v Local to any candidate steady state Z *consistent with (i)-(iv) there is an Anosov diffeomorphsim* $\mathcal{Z} \times U_t^{\mathcal{S}} \to f \times U_{t+1}^{\mathcal{S}}$

The last assumption is the non-linear equivalent of the familiar rank condition of linear algebra and will be met if each variable Z_t contributes a single error. Part (i) is a non-negligible restriction on the error process $\{U_t^S\}$ just because $\mathbb{E}C^{-\sigma}$ exists for some $\sigma > 0$ does not imply $\mathbb{E}C$ exists but I do not explore it here. The penultimate restriction is an accessibility condition it will mean limiting distributions are defined on the whole space. This will rule out hysteresis and multiple equilibria. It is uniformly adopted in informal formulations of the benchmark New Keynesian model.

Proposition 1. *Abstracting from policy shocks the Calvo model can be characterized by a canonical form with* $Z_t = (\pi_{t-1}, \pi_t, Y_t, \Delta_t)$ *and*
$U_t^S = (\phi_{t-1}, \phi_t, \varphi_{t-1}, \varphi_t, A_{t-1}, A_t)$

Proof. The main complexity arises with the derivation of the non-linear Phillips curve. This is because inflation via the reset price depends on the future of output, the discount factor and marginal costs - in a non-recursive fashion through $\aleph$ and $\beth$. Note that I have drawn equivalence between inflation and the reset price since the relationship is a diffeomorphism via the inverse function theorem since

$$\frac{d\pi_t}{d(p_t^*/P_t)} = \frac{(1-\alpha)}{\alpha}\left(\frac{p_t^*}{P_t}\right)^{-\theta}(1+\pi_t)^{2-\theta} > 0$$

The proof consists of using the lagged inflation relationship to eliminate the denominator $\beth_t$ from the present inflation relationship and solve for $\aleph_t$ in terms of present inflation π_t and then using the recursion of $\aleph_t$ to provide the recursion for inflation that forms the Phillips curve. Identical results would appear if I eliminated $\aleph_t$ and used the recursion of $\beth_t$, the crucial point is that I need to use a lagged inflation relationship in addition to the present inflation relation to create a simultaneous equation system that allows me to solve out for the two summation terms $\aleph_t$ and $\beth_t$. It will be essential to use inverse function arguments admitted by the differentiability assumptions in place, as non-linear stochastic systems rarely permit closed form expressions.

$$h(\pi_{t-1}) = \frac{u_c(C_{t-1})Y_{t-1}MC_{t-1} + \alpha\beta(1+\pi_t)^{\theta}\aleph_t + \tilde{f}_0(u_t^j)}{u_c(C_{t-1})Y_{t-1} + \alpha\beta(1+\pi_t)^{\theta-1}\beth_t + \tilde{f}_1(u_t^j)}$$

where the errors are bounded below by the positivity of the expectation terms $\mathbb{E}_{t-1}(1+$

$\pi_t)^\theta\aleph_t$ and $\mathbb{E}_{t-1}(1+\pi_t)^{\theta-1}\beth_t$ which follows from the non-negativity of expectations of positive random variables proved in the course of lemma 2.

$$\tilde{f}_0(u_t^j) > -\alpha\beta(1+\pi_t)^\theta\aleph_t$$
$$\tilde{f}_1(u_t^j) > -\alpha\beta(1+\pi_t)^{\theta-1}\beth_t$$

It is convenient to simplify the model at this point. Note that from the cost minimization and labor supply conditions we know that

$$u_c(C_t)MC_t = \frac{\nu_l(l_t)}{A_t}$$

Now using the market clearing conditions (12) and (24) note that

$$u_c(C_{t-1})MC_{t-1} = \frac{\nu'(\Delta_{t-1}Y_{t-1}/A_{t-1})}{A_{t-1}}$$

To obtain the required form it is necessary to remove Y_{t-1} and Δ_{t-1}. First writing the lagged Euler as

$$u'(Y_{t-1}) - \beta R_{t-1}(Y_{t-1},\pi_{t-1})\mathbb{E}_{t-1}\frac{u'(Y_t)}{1+\pi_t}\frac{\phi_t}{\phi_{t-1}} = 0$$

By the implicit function theorem I can solve out for $Y_{t-1} = \tilde{f}_3(\pi_{t-1},\pi_t,Y_t,\phi_{t-1},u_t^j)$ with $\tilde{f}_{3,1} < 0$, $\tilde{f}_{3,2} > 0$, $\tilde{f}_{3,3} > 0$ and $\tilde{f}_{3,4} > 0$. Note that I have used the properties of the Taylor rule discussed in Section 2.4. I can do likewise for lagged price dispersion using the stochastic analog of (40) shown in remark 2 $\Delta_{t-1} = \tilde{f}_4(\pi_t, u_{t-1}^\Delta)$ with $\tilde{f}_{4,1} < 0$ and $\tilde{f}_{4,2} < 0$.

$$h(\pi_{t-1}) = \frac{\tilde{f}_5(\pi_{t-1},\pi_t,Y_t,\Delta_t,\phi_{t-1},A_{t-1},u_{t-1}^\Delta,u_t^j) + \alpha\beta(1+\pi_t)^\theta\aleph_t}{\tilde{f}_6(\pi_{t-1},\pi_t,Y_t,\phi_{t-1},u_t^j) + \alpha\beta(1+\pi_t)^{\theta-1}\beth_t}$$

where $\tilde{f}_5$ has absorbed $\tilde{f}_0$ and likewise $\tilde{f}_6$ has absorbed $\tilde{f}_1$, both must be strictly positive by the functional restrictions on the errors and expectations laid out above, now solving for $\beth_t$ yields:

$$\beth_t = \frac{\tilde{f}_5 - h(\pi_{t-1})\tilde{f}_6 + \alpha\beta(1+\pi_t)^\theta\aleph_t}{\alpha\beta(1+\pi_t)^{\theta-1}h(\pi_{t-1})}$$

Note that since $\daleth_t$ is strictly positive by construction (34) and since the denominator is strictly positive the numerator must also be strictly positive. Hence, I can substitute into the present period reset price relationship

$$h(\pi_t) = \frac{\alpha\beta(1+\pi_t)^{\theta-1}h(\pi_{t-1})\aleph_t}{\tilde{f}_5 - h(\pi_{t-1})\tilde{f}_6 + \alpha\beta(1+\pi_t)^{\theta}\aleph_t}$$

Solving for $\aleph_t$ yields

$$\aleph_t = \frac{h(\pi_t)}{\alpha\beta(1+\pi_t)^{\theta-1}} \frac{\tilde{f}_5 - h(\pi_{t-1})\tilde{f}_6}{h(\pi_{t-1}) - h(\pi_t)(1+\pi_t)}$$

We are almost done however, there is the issue of a singularity to contend with. This corresponds to the case where $p_t^* = p_{t-1}^*$. Fortunately the singularity is removable since by manipulation of the reset prices I am able to solve for $\aleph_t$ for any initial $\aleph_{t-1}$ the expression is as follows

$$\aleph_t = \frac{\aleph_{t-1}}{(1+\alpha-\alpha(1+\pi_{t-1}))^{1/(\theta-1)}}$$

which lies strictly between $1/(1+\alpha)^{\theta-1}$ and $(1/\alpha)^{\theta-1}$ thanks to the lower and upper limits on inflation. Moreover, by composition of smooth functions - smoothness is retained. This position of the system corresponds to the stochastic steady states of a zero (trend) inflation economy. To complete the proof substitute into the recursion for $\aleph$ (35). The form is highly non-linear so solving for a closed form is inconvenient.Note that we can form an implicit function

$$\tilde{f}_7(\pi_{t+1}, Y_{t+1}, \Delta_{t+1}, \tilde{f}_5(.), \cdots, u_t^j) = 0$$

. Where the first three terms have entered through $\mathbb{E}_t(1+\pi)^{\theta}\aleph_{t+1}$. The appropriate vector of jump errors comes about by (ii) (v) and remark 2. Thus yielding the correct desired dimensionality $Z_t = (\pi_{t-1}, \pi_t, Y_t, \Delta_t)'$ and $U_t^S = (\phi_{t-1}, A_{t-1}, u_{t-1}^{\Delta}, \phi_t, A_t, u_t^{\Delta})$. To complete the derivation of the Phillips curve note that $\tilde{f}_7(Z_t, U_t^S)$ can be written as a composition of continuously differentiable functions that is locally non-constant in every variable it is a corollary of the mean value theorem that its set of stationary points will be measure zero. Therefore, it can be inverted almost everywhere for Z_{t+1}

using the implicit function theorem yields a function $\tilde{f}_8 : (Z_t, Y_{t+1}, \Delta_{t+1}, U_t^{\mathcal{S}}) \to \pi_{t+1}$ where the expectation exists by assumption 1. This topological justification will be used in all subsequent steps.

Finally inverting the Euler for Y_{t+1} in a fashion analogous to the construction of f_3 and substituting in the price dispersion relationship for π_{t+1} yields $\tilde{f}_9 : (Z_t, \pi_{t+1}, U_t^{\mathcal{S}}) \to \pi_{t+1}$ finally invert $\tilde{f}_9$ and then integrating over $U_t^{\mathcal{S}}$ gives $\mathbb{E}_t\pi_{t+1} = f_\pi(Z_t, U_t^{\mathcal{S}})$ to complete the derivation substitute in $\tilde{f}_9$ in to the Euler equation and integrate over the jump errors to solve for $\mathbb{E}_t Y_{t+1} = f_y(Z_t, U_t^{\mathcal{S}})$ and likewise for $\mathbb{E}_t\Delta_{t+1} = f_\Delta$ then note $\mathbb{E}_t\pi_{t-1} = \pi_t$ hence we have the appropriate structural form $\mathbb{E}_t Z_{t+1} = f(Z_t, U_t^{\mathcal{S}})$ that is valid *almost everywhere*. By assumption 1 (ii) it must apply everywhere.

For proof. Suppose the converse that there were some other variable Z^* such that $\mathbb{E}_t Z_t = f^*(Z_t, Z_t^*, U_t^{\mathcal{S}})$ such that $\mathbb{E}_t Z_{t+1} = f(Z_t^*, .)$ were a constant for μ almost every $(Z_t, U_t^{\mathcal{S}})$ but non-constant for some set $A_{\mathcal{Z}\times\mathcal{U}}$ with $\mu(A_{Z\times U}) = 0$ then a contradiction would arise for there will be a jump discontinuity of f^* at,at least one point (Z_0, U_0, Z_0^*) where $f(Z^*, .)$ was behaving as a non-constant function. From a topological perspective, for some open set $A_{\mathcal{Z}}$ the inverse image of $\mathbb{E}_t Z_{t+1}$ must contain subsets of the form $f^{-1}(A_{\mathcal{Z}}) = A'_{\mathcal{Z}} \cup X$ for some disjoint pair where $A_{\mathcal{Z}}$ is an open interval and X has measure zero. However, this means that X cannot be an open set of a Euclidean and therefore nor can $f^{-1}(A_{\mathcal{Z}})$. This proves f^* must have a discontinuity. Now consider the candidate state space form containing the additional variable Z^* that takes the form $Z_t = (\pi_{t-1}, \pi_t, Y_t, \Delta_t, Z_t^*)$ and $f = (\tilde{f}, f^*)$ where f^* is the recursion for Z^* such that $\mathbb{E}_t Z_{t+1}^* = f^*(Z_t, U_t^{\mathcal{S}})$. Then f inherit the discontinuity from f^* which completes the contradiction and shows the representation of the structural form is globally valid. $\square$

2.6 Forward-Solution Phillips Curve ZINSS

Analysis of the NK model begins with the recursive marginal cost Phillips curve. It describes how current inflation is determined by the present deviation of real marginal costs from steady state and the expectation of next period's inflation. It is derived from combining the optimal price-setting condition (15) and the price level construction equation (16) and log-linearizing

$$\pi_t = \kappa \hat{mc}_t + \beta \mathbb{E}_t \pi_{t+1} \tag{2.44}$$

Where $\kappa = (1-\alpha)(1-\alpha\beta)/\alpha$ the corresponding infinite horizon forward solution for inflation in terms of future marginal costs is

$$\pi_t = \kappa \sum_{i=0}^{\infty} \beta^i \mathbb{E}_t \hat{mc}_{t+i} \tag{2.45}$$

First linearize the marginal cost function (12)

$$\hat{mc}_t = \hat{w}_t - \hat{a}_t \tag{2.46}$$

Next linearize the production function

$$\hat{a}_t = \hat{y}_t - \hat{l}_t + \hat{\Delta}_t \tag{2.47}$$

With the assumption of a non-stochastic steady state we know that $\hat{\Delta}_t = 0$ now using optimal labor supply condition (8) we find

$$\hat{mc}_t = (\eta + \sigma)[\hat{y}_t - \frac{1+\eta}{\eta+\sigma}\hat{a}_t] \tag{2.48}$$

Now the current New Keynesian model uses the *efficient output gap* y_t^e defined as the log-difference between actual output $\hat{y}_t$ and the flexible price equilibrium output $\hat{y}_t^f$

$$y_t^e = \hat{y}_t - \hat{y}_t^f \tag{2.49}$$

Note that flexible price equilibrium output is given by

$$\hat{y}_t^f = \frac{1+\eta}{\eta+\sigma}\hat{a}_t \tag{2.50}$$

Crucially, the productivity term $\hat{a}_t$ cancels from the expression (32) that connects the desired notion of the output gap y_t^e to the deviation of marginal costs $\hat{mc}_t$

$$\hat{mc}_t = (\eta + \sigma) y_t^e \tag{2.51}$$

This yields the conventional recursive Phillips curve where inflation is a function of the efficient output gap and next period's expected inflation

$$\pi_t = \omega y_t^e + \beta \mathbb{E}_t \pi_{t+1} \tag{2.52}$$

Where the form of the composite parameter is

$$\omega = (\sigma + \eta) \frac{(1-\alpha)(1-\beta\alpha)}{\alpha} \tag{2.53}$$

Its forward solution is

$$\pi_t = \omega \sum_{i=0}^{\infty} \beta^i \mathbb{E}_t y_{t+i}^e \tag{2.54}$$

Accepting that we cannot have a perfect fit to test the model we need error terms in each equation therefore the final system is the following familiar Euler, Taylor and Phillips curve triplet.

$$y_t^e = \mathbb{E}_t y_{t+1}^e - \frac{1}{\sigma}(i_t - \bar{r} - \mathbb{E}_t \pi_{t+1}) + u_t^1 \tag{2.55}$$

$$i_t = \bar{r} + a_\pi \pi_t + a_y y_t^e + u_t^2 \tag{2.56}$$

$$\pi_t = \omega y_t^e + \beta \mathbb{E}_t \pi_{t+1} + u_t^3 \tag{2.57}$$

To make the calculation of the forward solution easier I follow the convention of eliminating the policy rule to give the matrix system:

$$\mathbb{E}_t X_{t+1} = \begin{bmatrix} \pi_{t+1} \\ y_{t+1}^e \end{bmatrix} = A X_t + B u_t \tag{2.58}$$

Where $X_t = \begin{bmatrix} \pi_t \\ y_t^e \end{bmatrix}$ $A = \begin{bmatrix} \beta^{-1} & -\omega\beta^{-1} \\ \sigma^{-1}(a_\pi - \beta^{-1}) & 1 + \sigma^{-1}(a_y + \omega\beta^{-1}) \end{bmatrix}$
$B = \begin{bmatrix} 0 & 0 & 1 \\ -1 & \sigma^{-1} & \sigma^{-1}\beta^{-1} \end{bmatrix}$ $(u_t)' = \begin{bmatrix} u_t^1 & u_t^2 & u_t^3 \end{bmatrix}$
From the general solution

$$X_t = \mathbb{E}_t \sum_{i=0}^{\infty} -A^{-(1+i)} B u_{t+i} \tag{2.59}$$

The three variables can be expressed as sums of the expected future shock terms.

$$\pi_t = \mathbb{E}_t \sum_{k=0}^{\infty} \zeta_\pi^1 u_{t+k}^1 + \zeta_\pi^2 u_{t+k}^2 + \zeta_\pi^3 u_{t+k}^3 \tag{2.60}$$

$$y_t = \mathbb{E}_t \sum_{k=0}^{\infty} \zeta_y^1 u_{t+k}^1 + \zeta_y^2 u_{t+k}^2 + \zeta_y^3 u_{t+k}^3 \tag{2.61}$$

$$i_t = \bar{\imath} + \mathbb{E}_t \sum_{k=0}^{\infty} \zeta_i^1 u_{t+k}^1 + \zeta_i^2 u_{t+k}^2 + \zeta_i^3 u_{t+k}^3 \tag{2.62}$$

The details of the ζ coefficients are not important here and are reported in Appendix 7.1. The Blanchard-Kahn condition (Blanchard and Kahn [1980]) that both eigenvalues of the matrix A lie outside the unit circle is required for the series to converge to a unique solution.

2.6.1 Persistence Problem

The persistence problem of the New Keynesian model lies in the property of the errors and expectations. To link the two I impose imperfect information.

Assumption 2. *The Central Bank has imperfect information about the present state of the economy - which is resolved at the end of the period after they have chosen their behavior.*

This restriction seems realistic for example quarterly output figures are released soon after the end of the relevant quarter, so the Central Bank only observes present state of the economy with error, this allows us to interpret the error term in the Taylor rule u_t^2 as a central bank expectation error.[14]

[14]In reality private sector agents should be equally if not more uncertain about the state of the economy. However, we do not need to assume imperfect information to interpret error terms u_t^1 and

Remark 2. *For (39)-(41) to constitute an identified rational expectations model each error term u_t^i must be white noise i.e.* $\mathbb{E}\, u_t^i \,|\mathcal{I}_{t-l} = 0,\ \forall\, t > 0,\ l \le t$

Here $\mathcal{I}_T$ is the information set provided by the model at time T or else the model is observationally equivalent to a model with bounded rationality (irrational expectations) and is therefore not identified in the encompassing class of DSGE models. This result follows from the fact that when we estimate the model with macroeconomic data we cannot observe expectations. Therefore we must use the mathematical expectation given by the structural model denoted by a superscript M. Take the example of the Phillips curve for concreteness $\mathbb{E}_t\pi_{t+1} = \mathbb{E}_t^M \pi_{t+1}$. Assumption 1 makes this easily applicable to both other equations in the system.Using S to denote 'subjective' and assuming there are no other I can subsume the Phillips curve error term u_t^3 into the subjective expectation yields $\mathbb{E}_t^S\pi_{t+1} = \mathbb{E}_t^M\pi_{t+1} + u_t^3$. Now if u_t^3 is not white noise then the agent is making systematically incorrect predictions and we have observational equivalence with a bounded rationality (irrational expectations) model. Therefore the rational expectations model is not identified[15].

For this reason apart from the technology shock, that as shown above does not feature in the final solution of the benchmark New Keynesian model the errors will be independently and identically distributed - consistent with the structural form.The following result characterizes the persistence problem.

Proposition 2. *There exists a solution to the linear equation system (39)-(41) in which all three major macroeconomic variables (y_t^e, i_t, π_t) as serially uncorrelated processes. If all the inverse eignevalues of the matrix A lie outside the unit circle this solution is unique.*

Proof. For illustration consider π_t analogous arguments can be made for the other two variables. As the system is ergodic the autocovariance generating function is symmetric so $Cov(\pi_t, \pi_{t-l}) = Cov(\pi_t, \pi_{t+l})\ \forall l$ so it suffices to show that $Cov(\pi_t, \pi_{t-l}) = 0$ for arbitrary lag length. From the white noise error assumption we know that $\mathbb{E}_t u_{t+k}^{\pi}|\,\mathcal{I}_t = 0\ \forall\, k > 0$ this means that current inflation can be expressed as a function

u_t^3 as expectation errors since these equations already contain the unknown variable $\mathbb{E}_t\pi_{t+1}$. Indeed introducing imperfect information to the private sector would require the unnecessary complication of a signal extraction problem.

[15] An alternative strategy would be to relax rational expectations and use survey data to measure inflation expectations Roberts [1995]Roberts [1997]Coibion [2010], although I do not pursue this here

of only contemporaneous shocks $\pi_t = \zeta_\pi^1 u_t^1 + \zeta_\pi^2 u_t^2 + \zeta_\pi^3 u_t^3$. Applying the same argument to period $t-l$ implies π_{t-l} is a function of time $t-l$ errors. Finally the white noise assumption means that $\forall\, i$ and $\forall\, l > 0$ $\mathbb{E}(u_{t-l}^i u_t^i) = 0$ so $Cov(u_{t-l}^i, u_t^i) = 0$. Now the result follows from noting that $Cov(\pi_t, \pi_{t-l}) = \sum_{i=0}^3 \sum_{j=0}^3 Cov(u_t^i, u_{t-k}^j)$ where every term in the summation is zero. $\square$

The no persistence result arises because the model lacks either intrinsic or extrinsic persistence. It lacks intrinsic persistence because the current value of the state variables X_t can be written as a function of just the current and future values of the shock process u_t independent of their past realizations.[16] It lacks extrinsic persistence because the errors are observationally equivalent to expectation errors which means they cannot be persistent. It will carry over to forward-looking policy rules that contain future expected inflation and output gap terms because they can be collapsed into the contemporaneous form (33) as the expectation of future variables will all be zero although verifying the existence condition on the eigenvalues will likely require a numerical computation routine.

The absence of persistence in the New Keynesian model contrasts with RBC models where the relevant solution variable is simply output y_t- which can inherit persistence from technology or news about future productivity developments Kydland and Prescott [1982] Beaudry and Portier [2006] Beaudry and Portier [2007] Jaimovich and Rebelo [2009] Walker and Leeper [2011] Schmitt-Grohé and Uribe [2012b]. Output itself inherits this persistence, it moves one-for-one with the efficient output in the corresponding RBC model so that the output gap stays constant. This highlights the point that Neo-Classical variables and associated shocks do not necessarily appear in the New Keynesian solution for business cycle dynamics. Therefore, novel New Keynesian features are needed to improve the fit of the New Keynesian model. It is for this reason that this paper introduces a new feature first-order price dispersion, unique to environments with price rigidity, to the basic New Keynesian model.

The econometric implications are unfortunate. The model cannot be used for forecasting. The key macroeconomic policy variable inflation is white noise. Neither can the model contribute to forecasting output or interest rates, over and above purely

[16] In the frequency domain this corresponds to the model acting as a neutral filter preserving the correlation spectrum in the error terms- a point first made by Cogley and Nason [1995b].

statistical procedures or classical models of the natural rate. To confirm: the New Keynesian model is not yet useful for policy. This is particularly unfortunate as the ability to forecast short-term fluctuations is the yardstick against which New Keynesian macroeconomists ask to be judged. In the article's conclusion (p87) the authors make their point strongly rejecting a suite of Real Business Cycle models on the following grounds: "We have demonstrated that the forecastable movements in output, consumption and hours [the three main variables in the Real Business cycle framework]"- what we would argue is the essence of the 'business cycle'- are inconsistent with a standard growth model disturbed solely by random shocks to the rate of technical progress." Rotemberg and Woodford [1996] (p71)[17] New Keynesian economics is not living by its professed econometric standards.
For those with sufficient perspective this is all rather reminiscent of the evolution of Classical economists' attitudes towards econometrics encapsulated in the following quote by Nobel laureate Thomas Sargent about fellow laureates Edward C (Ed) Prescott and Robert (Bob) Lucas: "My recollection is that Bob Lucas and Ed Prescott were initially very enthusiastic about rational expectations econometrics. After all, it simply involved imposing on ourselves the same high standards we had criticized the Keynesians for failing to live up to. But after five years of doing likelihood ratio tests on rational expectation models, I recall Bob Lucas and Ed Prescott both telling me that those tests were rejecting too many good models." Sargent [2005]

Worse still the basic model can not even be estimated with data from the three Q_t variables because it is not *identified* in the sense of Lehmann and Casella [1998] or Dufour and Hsiao [2008]. Note that to speak meaningfully of the *estimation* of a parameter $\mathbf{b}$ its set of possible values (parameter space) $\mathbf{B}$ must contain distinct values $\mathbf{b}_1 \neq \mathbf{b}_2$ as a single element space would correspond to *calibration*

Proposition 3. *The structural parameters of the model defined by (33)-(35) cannot be identified by frequentist econometric estimation.*

Proof. Applying proposition 1 simplifies the system to:

$$y_t^e = \frac{1}{\sigma}(i_t - \bar{r}) + u_t^1$$

[17] see also the abstract of Blinder and Fischer [1981] for a similar definition

$$i_t = \bar{r} + a_\pi \pi_t + a_y y_t^e + u_t^2$$

$$\pi_t = \omega y_t^e + u_t^3$$

Define the vector of endogenous variables $Q_t = (y_t^e, i_t, \pi_t)$ and the parameter vector $\theta = (\gamma, a_\pi, a_y, \omega, \beta, \lambda)$ where $\gamma = \frac{1}{\sigma}$ and λ is the collection of parameters governing the joint distribution of the three error terms. Θ denotes the sample space of the parameters formed of the product space $(\Gamma \times A_\pi \times A_y \times \Omega \times B \times \Lambda)$ where for example A_π is the set of admissible values for the parameter a_π. In the common case of normally distributed errors and unconstrained optimization $\Theta = \Re^{11}$ with Λ consisting of the distinct terms of the variance-covariance matrix of the error terms.

The crucial object is $f_{\theta_0}(Q_t)$ the joint probability distribution induced by the particular parameter vector $\theta_0 \in \Theta$ at time t. Recall that a parameter θ is identified when there is a one-to-one mapping to the probability distribution $\theta \to f_\theta(Q_t)$ at every time t.

Suppose the model were identified and proceed by the counterexample. Since β does not appear in the reduced form I can construct a counterexample with any $(\gamma, a_\pi, a_y, \omega, \lambda) \in (\Gamma \times A_\pi \times A_y \times \Omega \times \Lambda)$ and $\beta_1, \beta_2 \in B$ with $\beta_1 \neq \beta_2$ let $\theta_1 = (\gamma, a_\pi, a_y, \omega, \beta_1, \lambda)$ and $\theta_2 = (\gamma, a_\pi, a_y, \omega, \beta_2, \lambda)$ as $f_{\theta_1} = f_{\theta_2}$ $\forall t$ but $\theta_1 \neq \theta_2$ contradicting the hypothesis of a one-to-one mapping. □

For concreteness consider the popular Generalized Method of Moments Estimator [18] for the structural parameter vector $R_{3\times 3}$ of the relationship between Q_t and $\mathbb{E}_t Q_{t+1}$. It is necessary to have an $m \times 1$ $m > 3$ vector of available instruments $Z_t \in \mathcal{I}_t$ consistent with the orthogonality condition and associated estimator:

$$\mathbb{E}_t[Z_t'(Q_t - Q_{t+1}R)] = 0$$

To allow for the case of over-identification where the number of potential instruments in Z_t exceeds the number of moment conditions $m > 3$ in the basic model I minimize the quadratic form of the orthogonality conditions $H_T(R) = [T^{-1}(Z_t'(Q_t - Q_{t+1}R))W_T(Z_t'(Q_t - Q_{t+1}R))']$. Here W_T is a weighting matrix dependent on θ that

[18] The approach was developed by Hansen [1982] applied to rational expectations modeling by Hansen and Singleton [1982] and based upon method of moments estimation procedure first employed by Karl Pearson.

will turn out to be inversely proportional to the variance-covariance matrix of the orthogonality conditions, as previously T is the number of time observations. Optimization with continuous differentiability in R yields the GMM estimator[19]

$$\hat{R} = (Q'_{t+1}Z_tW_TZ'_tQ_{t+1})^{-1}Q'_{t+1}Z_tW_TZ_tQ_t$$

with $\hat{R} = (Q'_{t+1}Z_t)^{-1}Q'_{t+1}Q_t$ corresponding to the just-identified case where the orthogonality conditions are solved exactly.

However, applying proposition 2 again shows that this estimator is not defined because $Q'_{t+1}Z_t = 0$ meaning that the first matrix is non-invertible. This follows because Q_{t+1} is comprised entirely of expectation errors which must be uncorrelated with any variable belonging to the information set at time t, $\mathcal{I}_t$ to which Z_t belongs. Hence there are no valid instruments for the expectations of future macroeconomic variables in this system. This implies the structural parameters θ are not identified. For proof suppose the converse that θ were identified (i.e. there were sufficient valid instruments) assumption 2 would bound the expected deviation from the orthogonality conditions which would be sufficient to invoke Hansen and Singleton [1982] to prove weak convergence (in probability) of the GMM estimator $\hat{\theta}_T \to \theta$. Now consider the solution for the reduced form parameters in terms of their structural counterparts.

$$R_{11} = \frac{1}{1+\gamma(a_y+a_\pi)} > 0$$

$$-\infty < R_{12} = \frac{\gamma(a_\pi\beta-1)}{1+\gamma(a_y+a_\pi)} < \infty$$

$$R_{13} = 0$$

$$R_{21} = \frac{a_y+a_\pi\omega}{1+\gamma(a_y+a_\pi)} > 0$$

$$R_{22} = 0$$

$$-\infty < R_{23} = \frac{a_y\gamma(a_\pi\beta-1)+a_\pi[\omega\gamma(a_\pi\beta-1)+\beta(1+\gamma(a_y+a_\pi\omega))]}{1+\gamma(a_y+a_\pi)} < \infty$$

$$R_{31} = \frac{\omega}{1+\gamma(a_y+a_\pi)} > 0$$

[19] Consult Hansen and Singleton [1982] or a textbook such as Hamilton for a full exposition

$$R_{32} = 0$$

$$-\infty < R_{33} = \frac{\omega\gamma(a_\pi\beta - 1) + \beta[1 + \gamma(a_y + a_\pi\omega)]}{1 + \gamma(a_y + a_\pi)} < \infty$$

Note that each reduced form parameter R_{ij} is a composite of continuous functions and is therefore a continuous function of the structural parameters θ Denote this function by $\digamma_{ij}$ so $\digamma_{ij} = R_{ij}$. Therefore by the continuous mapping theorem of Mann and Wald [1943] $\lim_{T\to\infty} \hat{\digamma}_{ij} = R_{ij}$ *in probability.* This would create a one-to-one mapping between reduced-form parameters and probability distributions over the observables Q_t via the probability limits of the reduced form, the probability limits of the structural parameters and the structural parameters themselves. Hence the reduced form parameters would be identified- a contradiction. Therefore the structural parameters must be unidentified.

The notion that the New Keynesian Phillips curve might be weakly identified is well established Woodford [1994], Mavroeidis [2004], Koop et al. [2013] and Canova and Sala [2009] provides a typology of identification problems where this case corresponds to *observational equivalence* which he illustrates with a similar example.

2.7 Alternative Models and Equivalences

This section considers two alternative benchmark models of nominal rigidity and compares them to the non-linear Calvo setup of 2.2 and the benchmark linearized Calvo model in 2.1 - 2.4. Novel results appear. Some models thought to be equivalent are not whilst others thought to behave differently are the same in a particular sense.

2.7.1 Lucas Model

Lucas [1972] is the original rational expectations monetary model. It has a different structure to New Keynesian models: markets are perfectly competitive, anticipated monetary shocks are impotent. Monetary non-neutrality arises because of imperfect information rather than sticky prices. However, by adding a money demand function I am able to derive surprising similarity to New Keynesian models with price rigidity.

There are a collection of price taking households[20]. Each produces a single good using just its own labor[21] The choice problem simplifies to a static labor-leisure trade off

$$u_t(i) = C_t(i) - \frac{l_t^{1+\eta}}{1+\eta} \tag{2.63}$$

As before there are constant returns to scale and common technology

$$y_t(i) = A_t l_t(i) \tag{2.64}$$

Hence the firms objective function in $y_t(i)$ becomes

$$\frac{p_t(i)}{P_t} y_t(i) - \frac{(y_t(i))^{1+\eta}}{1+\eta} \tag{2.65}$$

The first order condition balances marginal revenue and marginal cost so

$$\frac{p_t(i)}{P_t} - (y_t(i))^{\eta} = 0 \tag{2.66}$$

Solving and log-linearizing gives an expression for an individual firm's supply curve

$$\hat{y}_t(i) = \frac{1}{\eta}(\hat{p}_t(i) - \hat{p}_t) \tag{2.67}$$

For parsimony a reduced form aggregate demand curve replaces the Euler equation (7).

$$\hat{y}_t = \hat{m}_t - \hat{p}_t \tag{2.68}$$

Idiosyncratic taste shocks reflected by a coefficient $Z_t(i)$ in the individual firms demand schedule are essential to the Lucas model. I assume they are independent and identically distributed across producers and time, with tail restrictions such that they

[20] It is not common to have differentiated products and price taking behavior. For motivation imagine there were actually multiple households producing the same good and therefore competition between them or the possibility to set-up firms with a consequent free entry condition.

[21] This formulation comes from Romer [2012]. The idea is to avoid the households using the economy-wide labor market to deduce the aggregate price level. In the original paper Lucas used an overlapping generations structure where households produced in one period then sold output in the next. Recent literature has developed more sophisticated stories about dispersed private information and strategic interaction. I do not pursue this approach as I am looking for a benchmark model.

will cancel out in aggregate almost surely. Here is the log-linear form

$$\hat{y}_t(i) = \hat{y}_t + \hat{z}_t(i) - \theta(\hat{p}_t(i) - \hat{p}_t) \tag{2.69}$$

Hence we can obtain the following approximate aggregation results

$$\hat{p}_t = \bar{\hat{p}}_t(i) \tag{2.70}$$

$$\hat{y}_t = \bar{\hat{y}}_t(i) \tag{2.71}$$

From these aggregations and the absence of other sources of inefficiency the first theorem of welfare economics tells us

$$\hat{y}_t = y_t^e \tag{2.72}$$

By steady state assumption[22]

$$\pi_t = \hat{p}_t \tag{2.73}$$

Monetary non-neutrality is entirely driven by unexpected shocks. There is no role for price dispersion here since the firms and price takers and the distribution of relative price is unaffected by inflation. Unlike its Keynesian counterparts elsewhere in this section the Lucas model would be unaffected by trend inflation. The driving force for this model is an inability for producers to distinguish aggregate to the price level P_t from idiosyncratic shocks $Z_t(i)$. Instead it only observes the price of its own good $p_t(i)$. A decomposition of the individual price change into real and nominal factors is instructive

$$\hat{p}_t(i) = \hat{p}_t + (\hat{p}_t(i) - \hat{p}_t) \equiv \pi_t + r_t(i) \tag{2.74}$$

where $r_t(i) = \hat{p}_t(i) - \hat{p}_t$ is relative price inflation - the change in the relative price of good i from period $t-1$ to the current period t. Thus the producer observes not current inflation but the sum of relative price change and general inflation.

[22] With this zero inflation steady state assumptions in place this model is equivalent to the earlier New Classical Phillips curve formulations from Phelps [1968] which lacks explicit microfoundations and feature non-rational expectations. It is also equivalent toFriedman [1968a] if initial expectations are correct $\mathbb{E}_{t-2}\pi_{t-1} = \pi_{t-1}$. This is of course not true in general. In particular with non-rational expectations predictable changes in inflation will effect output in so far as subjective expectations differ from rational expectations. This underlines the centrality of rational expectations to Lucas' model. However, under aggregation assumption (52) steady state conditions will be unaffected.

There is a *signal extraction* problem here with full information or absent uncertainty about either monetary policy or consumer preferences. The firm would prefer to base its production decision on relative prices alone. However, the producer does not observe $\hat{r}_t(i)$ but must estimate it given the observable own price $p_t(i)$. [23] Hence (49) becomes

$$\hat{y}_t(i) = \frac{1}{\eta}\mathbb{E}_t[r_t(i) \mid \hat{p}_t(i)] \tag{2.75}$$

Finally Lucas imposed functional form restrictions on the monetary shock $\hat{m}$ and $\hat{z}_i$ they are normally distributed with mean 0 and variances V_m and V_z and they are independent. The solution is obtained by guessing and verifying that π and r_i are normally distributed and independent. The hypothesis yields

$$\begin{aligned}\mathbb{E}_t[r_t(i) \mid \hat{p}_t(i)] &= \mathbb{E}_t[r_t(i)] + \frac{V_r}{V_r + V_\pi}(\hat{p}_t(i) - \mathbb{E}_t(\hat{p}_t(i))) \\ &= \frac{V_r}{V_r + V_\pi}(\hat{p}_t(i) - \mathbb{E}_t(\hat{p}_t(i)))\end{aligned} \tag{2.76}$$

where V_r and V_π are the strictly positive variances of relative prices and inflation respectively. Note that the ratio of V_r to V_π is the *signal- to-noise ratio.* Finally substituting (58) into (57) and then aggregating with (53) yields the famous Lucas supply curve

$$\hat{y}_t = b(\pi - \mathbb{E}_t\pi_t) \tag{2.77}$$

Where

$$b = \frac{1}{\eta}\frac{V_r}{V_r + V_\pi} \tag{2.78}$$

To finish it is necessary to solve out for V_r and V_π in terms of the structural parameters of the model V_m and V_z the approach is as follows- solve for π and y^e using aggregate

[23] Recall that the firm is owned by a single household. If the household knew others' prices through making purchases, it could deduce $p_t(i)$ and hence $r_t(i)$. There are several ways to rule this out. The most common is to assume each household can be divided into a "producer" and a "shopper" who do not communicate. Alternatively with Lucas' original OLG approach the problem is avoided because the individual produces in the first period and shops in the second. Jonung et al. [1981], Huber [2011], Ashton et al. [2012], Del Missier et al. [2016] and Detmeister et al. [2016] provide direct empirical support for the signal extraction hypothesis. In particular behavioral bias coming from the accessibility heuristic (that individuals over-weight common items when estimating their personal inflation rate) can mimic signal extraction even when information about the aggregate price level is in fact easily available.

demand curve (50) with inflation (55) and Lucas supply curve (59) yields

$$\pi_t = \frac{1}{1+b}\hat{m}_t + \frac{b}{1+b}\mathbb{E}_t\pi_t \tag{2.79}$$

$$\hat{y}_t = \frac{b}{1+b}\hat{m}_t - \frac{b}{1+b}\mathbb{E}_t\pi_t \tag{2.80}$$

Passing expectations through (61) yields

$$\mathbb{E}_t\pi_t = \mathbb{E}_t\hat{m}_t \tag{2.81}$$

Using (63) and the fact that $\hat{m}_t = \mathbb{E}_t\hat{m}_t + (\hat{m}_t - \mathbb{E}_t\hat{m}_t)$ (61) and (62) can be rewritten as

$$\pi_t = \mathbb{E}_t\hat{m}_t + \frac{1}{1+b}(\hat{m}_t - \mathbb{E}_t\hat{m}_t) \tag{2.82}$$

$$\hat{y}_t = \frac{b}{1+b}(\hat{m}_t - \mathbb{E}_t\hat{m}_t) \tag{2.83}$$

Using the individual supply curve (49), demand curve (51), steady state deviation (55) and the Lucas supply curve (59),(64) implies the expression for relative price deviation

$$\hat{p}_t(i) - \hat{p}_t = \frac{\hat{z}_t(i)}{\theta + b} \tag{2.84}$$

Hence $V_r = \frac{V_z}{(\eta+b)^2}$ and from (64) $V_p = \frac{V_m}{(1+b)^2}$. This leads allows me to derive the following implicit formulation final form for the slope of the Lucas surprise Phillips curve

$$b = \frac{1}{\eta}\left[\frac{V_z}{V_z + \frac{(\theta+b)^2}{(1+b)^2}V_m}\right] \tag{2.85}$$

Lucas focused on the limiting case lim $\theta \to 1$ which yields $b = \frac{1}{\eta}\frac{V_z}{V_z+V_m}$ Hence the final form of the Phillips curve comparable to Calvo is

$$\begin{aligned}\pi_t &= b^{-1}\, y_t^e \\ \pi_t &= \eta\,(\frac{V_m}{V_z} + 1)\, y_t^e\end{aligned} \tag{2.86}$$

where I have employed aggregation (54). Although, the forthcoming proposition will concern the general form derived from the implicit expression for b^{-1} from (69).

The last step is to verify the conjectures made concerning π and $\hat{r}_i$ in decomposition equation (58). Equations (61) and (66) imply that both are linear functions of $\hat{m}$ and $\hat{z}_i$. Since $\hat{m}$ and $\hat{z}_i$ are independent, π and $\hat{r}_i$ are also mutually independent. Since linear functions of normal variables are normal, π and $\hat{r}_i$ are also normally distributed and the derivation is complete.

The internal logic of the model is communicated through equations (64) and (65) that determine the relationship between the two key macroeconomic variables the output gap y^e and inflation π and the policy instrument in this case the money supply (in deviation form) $\hat{m}$. The component of aggregate demand that is observed, $\mathbb{E}\,\hat{m}$ affects only prices passing straight through into inflation. However, the part that is unobserved $\hat{m} - \mathbb{E}_m$ has real effects. Consider a positive shock to the money supply $\hat{m}$. This increases aggregate demand- producing an outward shift to each individual producers demand curve. Since aggregate demand is not observed, each supplier's best guess is that some portion of the rise in product demand represents a shock (positive in this case) to individual demand through the relative price. Thus producers increase output.

The effect of an observed increase in $\hat{m}$ is completely different. Suppose there is an anticipated increase in aggregate demand so $\mathbb{E}\,\hat{m}$ is increased with $m - \mathbb{E}\,\hat{m}$ held constant. Now each producer attributes the rise in demand for their product solely to the expansion of the money supply and thus there is no change in aggregate supply. Therefore observed movements in aggregate demand affect only prices.

The policy implications of this model as pointed out by Sargent and Wallace [1975] and Barro [1976] are stark. Monetary policy understood as systematic movements in aggregate demand reflected here in $\mathbb{E}\,\hat{m}$ will have no effect on the path of real variables output summarized here $\hat{y}$ because agents with rational expectations will anticipate changes in demand and see through them. Only the unpredictable part of demand $m - \mathbb{E}\,m$ will matter for real output. However, this roll of a dice is not what is understood by monetary policy. In particular, counter-cyclical policies- the common sense that Central Banks should lean against the macroeconomic cycle- cutting rates during a downturn and raising them when there is an unsustainable boom- are ineffective. Policy ineffectiveness remains the most powerful result in monetary policy analysis. I will show that New Keynesian economics has not entirely disentangled itself from its web.

2.7.2 Rotemberg Pricing

This model Rotemberg [1982] is perhaps the simplest model of state dependent pricing and the one most commonly compared to Calvo[24]. The idea is that firms face a convex cost of changing prices and therefore they do not adjust immediately to their optimal flexible price. These costs need not be viewed as just physical costs of price changing, although these can be substantial for some retailers but could also include the cost of finding and processing information about the market and the aggregate economy, negotiation with suppliers, communication with the public or as a short cut to incorporate behavioral responses of customers to price changes or fairness concerns on the part of producers.[25][26] Levy et al. [1997], Zbaracki et al. [2004], Gorodnichenko and Weber [2016] and Noton [2016] provide empirical strategies to measure such costs. The household problem, technology and policy rule will be as before. The difference lies in the source of price rigidity. The firm faces a cost of changing prices usually parametized as a quadratic and scaled by c_p

$$C_t^a(i) = \frac{c_p}{2}\left(\frac{p_t(i)}{p_{t-1}(i)} - 1\right)^2 \tag{2.87}$$

Profit maximization requires firms:

$$\max_{\{p_T(i), l_T(i)\}} \mathbb{E}_t \sum_{T=t}^{\infty} Q_{t,T}\left[\frac{P_T(i)}{P_T} y_T(i) - W_T l_T(i) - \frac{c_p}{2}\left(\frac{p_t(i)}{p_{t-1}(i)} - 1\right)^2\right] \tag{2.88}$$

Note unlike with Calvo pricing the firms' problem is recursive in present discounted profits with

$$V_t(p_{t-1}(i)) = \frac{P_t(i)}{P_t} y_t(i) - W_t l_t(i) - \frac{c_p}{2}\left(\frac{p_t(i)}{p_{t-1}(i)} - 1\right)^2 + \mathbb{E}_t[Q_{t,t+1} V_{t+1}]$$

[24] Alternative menu cost models with fixed costs of price changes are also popular but benchmark formulations such as Mankiw [1985] and subsequent extensions such as Nakamura and Steinsson [2008] feature discontinuous adjustment and therefore cannot be rendered in a log-linear form.

[25] See Blinder et al. [1990], Blinder et al. [1991] and Bertola et al. [2012] for supportive survey evidence.

[26] Theoretical papers that take up these ideas include Hobijn et al. [2006], Reis [2006], Gorodnichenko [2008] Alvarez et al. [2014],Alvarez et al. [2015], Maćkowiak and Wiederholt [2015] and Alvarez et al. [2016] for information processing stories and Renner and Tyran [2004], Rotemberg [2005], Nakamura and Steinsson [2011] and Rotemberg et al. [2011] concerning customer markets and behavioral approaches.

The first order conditions for an individual firm is

$$(1-\theta)\left(\frac{p_t(i)}{P_t}\right)^{-\theta} - \; c_p\left(\frac{p_t(i)}{p_{t-1}(i)}-1\right)\frac{Y_t}{p_{t-1}(i)}$$
$$+\theta MC_t(i)\left(\frac{p_t(i)}{P_t}\right)^{-(\theta+1)}\frac{Y_t}{P_t}+\mathbb{E}_t Q_{t,t+1}c_p\left(\frac{p_{t+1}(i)}{p_t(i)}-1\right)\frac{Y_{t+1}p_{t+1}(i)}{p_t(i)}=0 \quad (2.89)$$

where the top line is the marginal revenue net of the cost of price changing and the second reflects the effect on present and future marginal costs of an adjustment today.

$$\frac{p_t(i)}{p_{t-1}(i)} = 1+\pi_t \quad (2.90)$$

$$\Delta = 1 \quad (2.91)$$

Finally the resource constraint reflecting a wedge between production and consumption caused by the cost of price changing replaces (9)

$$Y_t = C_t + c_p \pi_t^2 Y_t \quad (2.92)$$

Flexible price output is the same as in the Calvo model and at the non-stochastic zero inflation steady state $\hat{c}_t = \hat{y}_t$ to yield the Phillips curve:

$$\pi_t = \tilde{\omega} y_t^e + \beta \mathbb{E}_t \pi_{t+1} \quad (2.93)$$

Where

$$\tilde{\omega} = \frac{(\sigma+\eta)(\theta-1)}{c_p} \quad (2.94)$$

2.7.3 Equivalence Conditions and Policy Implications

The connection between the models is as follows.

Proposition 4. *Consider the Phillips curve slope parameterizations* $\{w, b^{-1}, \tilde{w}\}$ *for Calvo, Lucas and Rotemberg models defined by (36)-(37), (69)-(70) and (77)-(78) respectively along with the set of common parameters* $\gamma^c = \{\sigma, \beta, \eta, V_m\}$ *and model specific parameters* $\gamma^s = (\alpha, c_p, V_z)$ *There exist settings for the common parameters that make the Rotemberg and Calvo models observationally equivalent for for all positive*

Phillips curve slopes and for the Lucas model also provided that the slope parameter exceeds unity.

Proof. Write a general Phillips curve as $\pi_t = g y_t^e$ where g is the slope of the Phillips curve. For the Calvo model first of all apply proposition 1 to show that the expected inflation term will always be zero. Then consider the slope as a function of the reset parameter α note crucially this parameter is unique to the Calvo model. Note further $\lim_{\alpha\to 1}\omega(\alpha) = 0$ and $\lim_{\alpha\to 0}\omega(\alpha) = \infty$. As $\omega(\alpha)$ is a composition of power and linear functions it must be continuous. Hence, apply the intermediate value theorem to show that for every $g > 0$ there is a parametization γ^m such that $\omega = g$.
Likewise for Rotemberg note that c_p is its unique parameter with $\lim_{c_p\to\infty}\tilde{\omega} = 0$ and $\lim_{c_p\to 0}\tilde{\omega} = \infty$ and apply the intermediate value theorem for $\omega(c_p)$ to complete the proof that for every $g > 0$ there is a parametization γ^m such that $\omega = \tilde{\omega} = g$ and apply proposition 1- to complete the proof. Its implementation is $c_p = \frac{\alpha(\theta-1)}{(1-\alpha)(1-\alpha\beta)}$.
Turning to the Lucas model. The proposition follows if I can show that a unique b exists for every possible implicit form (69) and that $0 < b < 1$ with continuity in parameters (69) note that by the parameter restrictions the right hand side is strictly positive for all possible b hence $b > 0$ next note that $\theta > 1$ ensures both terms in the denominator are strictly positive for all possible b hence we can only have $0 < b < 1$. Also note this bound cannot be tightened since for all other parameter values $\lim b_{V_z\to 0} = 0$ and $\lim b_{\eta\to 1, V_z\to\infty} = 1$
The proof is completed by an implicit function theorem argument completed with the identical intermediate value theorem argument. Define

$$F = b - \frac{1}{\eta}\left[\frac{V_z}{V_z + \frac{(\theta+b)^2}{(1+b)^2}V_m}\right] \equiv b - \frac{1}{\eta}\frac{(1+b)^2 V_z}{(1+b)^2 V_z + (\theta+b)^2 V_m} = 0$$

and its partial derivatives can be characterized as follows

$$F_b = 1 - \frac{2}{\eta}\frac{(1+b)V_z}{(1+b)^2 V_z + (\theta+b)^2 V_m} + \frac{1}{\eta}(1+b)^2 V_z \frac{[2(1+b)V_z + 2(\theta+b)V_m]}{((1+b)^2 V_z + (\theta+b)^2 V_m)^2}$$

Now the task at hand is to give F_b a definitive sign converting to a common denominator and a little cancellation gives

$$\frac{\eta(1+b)^4V_z^2 + 2\eta(1+b)^2(\theta+b)^2V_zV_m + \eta(\theta+b)^4V_m^2 - 2(\theta-1)(1+b)(\theta+b)V_zV_m}{((1+b)^2V_z + (\theta+b)^2V_m)^2}$$

Finally combining the second and fourth terms yields a strictly positive expression

$$F_b = \frac{\eta(1+b)^4V_z^2 + 2\eta[(\eta-1+b)\theta + 1 + \eta(1+b)b](1+b)(\theta+b)V_zV_m + \eta(\theta+b)^4V_m^2}{((1+b)^2V_z + (\theta+b)^2V_m)^2}$$

This proves $F_b > 0$. The other partial derivatives can be characterized as follows $F_\eta = \frac{1}{\eta^2}V_z/(V_z + (\frac{\theta+b}{1+b})^2)^2 > 0$, $F_\theta = \frac{2}{\eta}\frac{(\theta+b)^2}{1+b}V_z/(V_z + (\frac{\theta+b}{1+b})^2V_m)^2 > 0$, $F_{V_m} = \frac{1}{\eta}(\frac{\theta+b}{1+b})^2V_z/(V_z + (\frac{\theta+b}{1+b})^2V_m)^2 > 0$ and $F_{V_z} = -\frac{1}{\eta}(\frac{\theta+b}{1+b})^2/(V_z + (\frac{\theta+b}{1+b})^2)^2 < 0$ all are non-zero as required by the implicit function theorem.

As a composition of smooth functions F is itself smooth. Hence by the implicit function theorem b exists is unique and continuously differentiable in parameters (η, θ, V_z, V_m). It is therefore continuous in these arguments, as is b^{-1} by preservation of continuity under function composition and the intermediate value theorem on $1 < b^{-1}(\eta, \theta, V_z) < \infty$ proves that as with Rotemberg and Calvo in the relevant range I can always find common parameters such that $b^{-1} = g$. Since I used different members of γ^s for each parametization- I am able to vary them independently. Hence they are observationally equivalent for every possible setting for the common parameters provided $g > 1$. $\square$

The interpretation is that provided the Phillips curve is sufficiently steep then the methods of macroeconometrics cannot be used to distinguish between these three underlying designs for the Phillips curve. V_m has been assigned to the common parameters because it would be easy to add a monetary sector to either of the New Keynesian models without affecting topological properties beyond dimensionality if we cared to add a money demand shock. Moreover, it can be estimated in principle from monetary aggregates in a model free fashion.[27] The model specific parameters

[27] In practice there are issues with how money is defined and measured. See Hendry and Ericsson [1991], Friedman and Schwartz [1991], Ericsson et al. [2016] and Lucas and Nicolini [2015] for a window into this debate. Note there are also issues to do with the construction and availability of data- for monetary aggregates.

are those that only feature in a single model and need to be estimated from microeconomic data. This accords with the conventional practice in macroeconometric model comparison of calibrating common parameters and estimating parameters specific to a single model.

Without microeconometric evidence the standard New Keynesian model is equivalent to the Rotemberg model which contradicts the findings of 2.4 and 2.7 concerning their non-linear forms. Furthermore, if there is sufficient price flexibility it is also equivalent to its New Classical counterpart. This should be a surprising result given that the express purpose of the New Keynesian framework was to differentiate itself from earlier Classical monetary theory. Moreover, the fact that there can be equivalence with a non-dynamic model is indicative that the benchmark New Keynesian model has a weak claim to be the pinnacle of Dynamic Stochastic General Equilibrium modeling. Intuitively, if one misses out on the 'stochastic' part of the equilibrium then one loses the 'dynamics' that ought to a define a successful Dynamic Stochastic General Equilibrium.

Chapter 3

The Role of Price Dispersion

This Chapter builds on Section 2.3. It begins by analyzing the global properties of Δ as an aggregator which will prove crucial to subsequent chapters. I generalize this notion across a collection of sticky price models then compare the notion of price dispersion in a Keynesian context to a leading counterpart from New Classical macroeconomics. The difference lies with rigidity of the aggregate price level caused by the absence of selection as to which firms change price. I consider dynamic properties specific to the Calvo setting and by way of example application to zero bound models. I link the condition that $\Delta \geq 1$ to underlying non-negativity constraints. I link coordination failure on the production side before discussing extensions that can incorporate idiosyncratic shocks and some state dependence.

3.1 Lower Bound on Price Dispersion

The following powerful result derived directly from the construction of the price level tells us that the measure of price dispersion Δ defined by equation (18) in Section 2 is strictly greater than unity unless all firms set the same price.

Proposition 5. $\Delta \geq 1$ *with* $\Delta = 1$ *if and only if* $p_t(i) = P_t$, $\forall\, i$.

Where $p_t(i)$ is the price set by any firm i at time t. The lengthy proof is contained in Appendix A, the first part is a familiar application of Jensen's inequality, possible because the demand system is sufficiently convex, and the second a small extension exploiting strict convexity. Versions of this result are well known in the literature- see for example Schmitt-Grohe et al. [2007] or Damjanovic and Nolan [2010], although full proofs are omitted and analysis is usually restricted to non-stochastic steady states. The extension to other New Keynesian models that use the constant elasticity of substitution preference scheme is straightforward. I do so in Appendix A by modifying the probability measure used to aggregate the various prices to obtain the price level to correspond to three common pricing models: the basic Calvo model used here, the Calvo model with indexation to trend inflation used by Yun [1996] and the General Taylor Economy of Taylor [1993b], Coenen et al. [2007], Dixon and Kara [2010], Dixon and Kara [2011] and Dixon and Le Bihan [2012] which encompasses a wide range of pricing models and can be fitted exactly to match cross-section price distributions. [1]

The economic force driving this result is preference for variety- where the individual prefers averages to extremes. It ensures that variations in price unrelated to marginal costs make the individual worse-off. This is a very weak economic assumption. A strict preference against variety would make it difficult to ensure interior demand system for all products. The only cases in widespread usage where this argument would not prove widely applicable are those with homogeneous products as with Bertrand, perfect or Cournet competition. Therefore the result is certainly not specific to one demand system.

To allow for heterogeneity between firms I can redefine Δ to normalize each firm's price relative to its optimal reset price.Yun and Levin [2011], Fuhrer [2000], Dennis [2009], Ravn et al. [2010], Givens [2013], Santoro et al. [2014], Lewis and Poilly [2012], Lewis and Stevens [2015] and Etro and Rossi [2015] consider various alternative demand systems with a variety of motivations. I could alter the source of the price dispersion from staggered to for example information-constrained price-setting. Mankiw and Reis [2002], Mankiw and Reis [2006], Mankiw and Reis [2007], Lorenzoni [2009], Lorenzoni [2010], Nimark [2008], Nimark [2014], Adam [2007], Mackowiak et al. [2008] and Paciello and Wiederholt [2014] are papers where this price dispersion is present but not accounted for. All that is required is a motivation for firms to set different prices when in a flexible price world it would be efficient if they all set the same.

[1] Dixon [2012] shows that the Generalized Taylor model can approximate arbitrarily well the Generalized Calvo used by authors such as Wolman [1999], Dotsey and King [2006], Sheedy [2010], the multiple Calvo associated with Carvalho [2006] and de Carvalho [2011], as well as the familiar simple Taylor and Calvo.

Price dispersion would also come about where there are physical costs of price changing provided that firms face idiosyncratic shocks or differing adjustment costs. See for example, Gertler and Leahy [2008], Nakamura and Steinsson [2008], Reiff et al. [2014], Bouakez et al. [2009] and Bouakez et al. [2014]. In these cases the relevant interpretation of the price dispersion variable Δ is the difference between the actual and flexible price. Suppose for example fixed adjustment costs of a price change varying across firms with an aggregate shock - $\Delta > 1$ will come about if some firms adjustment costs are below and some above the common adjustment threshold. Similarly, with common fixed adjustment cost but idiosyncratic shocks $\Delta > 1$ will occur if some firms keep their price constant because their idiosyncratic shock 'cancels out' the aggregate shock i.e. they remain inside their band of price inaction. The behavior of price dispersion and its dynamics are integral to all the analysis which follows. All results that do not refer to a specific specification of sticky price setting (e.g. Calvo or Taylor) generalize to all models covered by this lemma. The fundamental mechanism in this paper is that inflation causes price dispersion.

3.1.1 Different Pricing Models

This subsection formalizes generalizations to other pricing setups and in so doing explains why several competiitor models, two of which were laid out in Section 2.6 do not generate a dynamic variable analagous to Δ in the basic Calvo set of Section 2.4. To this end I consider a new concept-*Real Price Dispersion*- the price dispersion that would exist in a models *flexible price limit*- where all prices could be reset in every period with perfect information at no cost to the firm- as in the RBC framework and standard microeconomic analysis[2] Real price dispersion could be thought of as the price dispersion that is *efficient* from the point of view of firms assuming no restrictions or costs of price setting and taking the constraints of the market structure and production technology as given. It takes the form

$$\Delta_t^* = \int_i F_t(i)\left(\frac{p_t^{**}(i)}{P_t}\right) \mathrm{d}\mu_t(i) \tag{3.1}$$

[2]In New Monetarist models firms prices are perfectly flexible whilst consumers face a costly search process to ascertain the distribution of prices- see Williamson et al. [2010] and the original contribution from Burdett et al. [1983] which shows search costs can give rise to a non-degenerate distribution of prices.

where

$$P_t = \int_i p_t(i) F_t(i) \left(\frac{p_t^{**}(i)}{P_t} \right) \, \mathrm{d}\mu_t(i)$$

is the price level. $F(i)$ represents the demand curve for firm i and is homogeneous of degree zero in prices - as in basic consumer theory. This allows me to define *nominal price dispersion* as the ratio between *actual* and *real* price dispersion. This can be viewed as price dispersion that is viewed as inefficient from the point of view of a firm operating in a flexible price environment

$$\Delta_t = \frac{1}{\Delta_t^*} \int_i F_t(i) \left(\frac{p_i}{P_t} \right) \, \mathrm{d}\mu_t(i) \tag{3.2}$$

This definition is consistent with the parametric form of Δ_t in the Calvo model given in (2.39) where I assumed firms faced the same demand and technology with no idiosyncratic shocks, which makes any price dispersion inefficient.

As before i is a number used to index an individual firm. Since I am allowing heterogeniety among firms' optimal prices and the possibility for multiple equilibrium, it is necessary to be more precise about how i is assigned. i reflects the order of the firm in the price distribution. Therefore, there exists a positive monontonic relationship between i and $p_t(i)/P_t$. This also simplifies the existence of the defining integral. [3] $\Omega_t(i)$ is the set of all prices in the economy at time t. Σ_t is the family of sets of individual firms over which output can be aggregated. [4]

Hence $p_t^{**}(i)$ represents the price that a firm at the corresponding point in the price distribution would set in the limit where *all* prices became flexible and information perfect as in the RBC framework. This differs from p_t^* which features in staggered price setting models such as the Calvo model in Chapter 2 and Taylor contracting models in this Chapter, where p_t^* represents the price a firm with a price that is fully

[3] Formally I have defined that Σ_t contains a countable family of pure points representing firms interacting strategically together and another family of Borel sets which are continua of firms who take the aggregate economy as given corresponding to perfect or monopolistic competitions as in Calvo, Taylor and other New Keynesian models. The associated measure $\mu(i)$ is positive as it reflects shares of goods in aggregate consumption- the existence of a discrete Lebesgue integral over the pure point sets follows immediately. Royden and Fitzpatrick proves that monotone functions on Borel sets in $\mathbb{R}$ possess Lebesgue integrals. This demonstrates that a measure μ_t exists over all sets in Σ_t.

[4] In mathematical terms Ω_t is the measurable space of firms and Σ_t is the smallest sigma-algebra which contains sets of firms that produce positive output and the empty set with no firms in it. The distinction between membership of Σ_t and Ω_t is operative in the specification of many macroeconomic models with imperfect competition which use continuum of firms, for which countable subsets of prices would belong to Ω_t but not Σ_t. The two coincide in real life where the set of prices at a point in time is countable.

flexible today would set taking into account that other prices in the economy are rigid. The notion that rigidity of prices elsewhere in the economy causes flexibly reset prices to differ from those that would be set in a fully flexible world ($p_t^{**}(i) \neq p_t^*(i)$) is known as *real rigidity* see Ball and Romer [1990]. It is present in major New Keynesian models. The RBC skeleton has a symmetric equilibrium so $p_t^{**} = P_t$ however the optimal reset price p_t^* only equals P_t when $\pi_t = 0$ otherwise $p_t^{**} \neq p_t^*$ so there is real rigidity. It would be interesting to explore the link between these two concepts further. To complete the generalization it is necessary to define two properties that DSGE models or more specifically their associated price distributions may possess. The first is *aggregate nominal rigidity*

Definition 2. An economy possesses aggregate nominal rigidity if there exists a measurable set of firms $\mathcal{B}^1 \in \Sigma_t(i)$ such that for some $l > 0$, $p_t(i) = \Phi_l(i)(p_{t-l}(i))$ and there exists no more than one inflation rate $\bar{\pi}(\Sigma_t)$ such that $\pi_t = \bar{\pi}_t$ implies that $p_t(i) = p_t^{**}(i)$ for $i \in \mathcal{B}^1$. Also it must be the case that $\forall\, l'$ where $0 < l' < l$ we can write $p_{t-l'}(i) = \Phi_{l-l'}(i)(p_{t-l}(i))$ where $\pi_{t-l'} \neq \bar{\pi}(\Sigma_t)$ implies $p_{t-l'}(i) \neq p_{t-l'}^{**}(i)$

The first part states that to have aggregate price rigidity there must be a positive fraction of output sold at a price that reflects past prices l periods back and differs from those that would prevail in the flexible economy. There is an allowance that if inflation hits a certain value the two could coincide as would occur in Calvo contracting starting from no price dispersion when inflation is zero or equal to a target $\bar{\pi}$ to which all prices are indexed either directly as in Yun [1996] or to the last periods inflation as in Smets and Wouters [2003], Christiano et al. [2005] and Smets and Wouters [2007][5]. The second part serves to ensure dependence between past and current price levels and therefore past and current levels of price dispersion. This means there cannot be aggregate price rigidity in an otherwise flexible economy just because there are backward-looking or cycling prices. This restriction has economic content. Many items are sold on temporary discount which upon expiry return to their old level.

Fortunately, recent New Keynesian models that feature products on sale avoid this

[5] With Taylor contracts we have to allow for the possibility that there may exist reset prices consistent with no inefficient price dispersion that differ over time. This is because prices replace one another under Taylor contracting see equation (49) so the price that sets Δ_t to one will depend on the price it is replacing in that case p_{T-M}^*.

trap as they are able to generate changes in the frequency or size of discounts in response to monetary shocks that would be neutral in the model's RBC skeleton- which implies purchases are made at $p_t \neq p_t^{**}$ see Kehoe and Midrigin [2008], Guimaraes and Sheedy [2011], Nakamura and Steinsson [2011], Eichenbaum et al. [2011] and Malin et al. [2015].[6]

As well as traditional RBC models, the definition of aggregate price rigidity excludes recent New Monetarist models. In this framework there is a distribution of prices motivated by a flexible price microeconomic model, it has become common to use a model with costly search such as Burdett et al. [1983], Albrecht and Axell [1984] and Burdett and Mortensen [1998] as these can provide a rationale for agents to hold money if there are appropriate credit constraints see Lagos and Wright [2005] and Williamson et al. [2010]. Their point is that provided the distributions of prices overlap from period to period it is possible for some firms not to change their price in equilibrium. Even though money will be exactly neutral because the market equilibrium in their model does not depend on the money supply. They claim therefore that price rigidity does not imply monetary non-neutrality.

Their claim is justified precisely because their model does not possess aggregate nominal rigidity. It has equilibria where individual firms choose to keep prices fixed because they do not care about their position in the price distribution by equilibrium construction. On the other hand, with my approach of indexing firms by their position in the distribution there is no nominal rigidity. At each point in the distribution the appropriate firm raises their price one-for-one with the money supply- so the aggregate price level is perfectly flexible.

It is necessary to impose one further condition on the price distribution this is called *nominal heterogeniety*. This states that the degree of nominal distortion represented by the ratio between the actual price and the flexible model price $p_t(i)/p_t^{**}(i)$ must vary between firms.

Definition 3. An economy possesses aggregate nominal heterogeniety if $\exists\, \mathcal{B}^1,\ \mathcal{B}^2 \in \Sigma_t(i)$ such that $\int_{\mathcal{B}^2} p_t(i)/p_t^{**}(i)\ \mathrm{d}\mu(i) - \int_{\mathcal{B}^1} p_t(i)/p_t^{**}(i)\ \mathrm{d}\mu(i) > 0$ for all $\pi_t = \bar{\pi}_t$

[6] Note that as these papers do not tend to use log-linearization around a non-stochastic steady state to characterize business cycle dynamics, they might be immune to some of the criticisms involving price dispersion in this chapter and subsequent concerns about the qualitative behavior of dynamic approximation discussed in Chapter 5.

This restriction rules out stylized models such as those in Section 2.6 along with contemporaries such as Barro [1972], Sheshinski and Weiss [1977], and Mankiw [1985] where all firms set the same price motivated by physical costs to price changing that do not differ among firms, as argued earlier such models are often unable to generate sufficient nominal rigidity. In any case physical costs of price changing vary substantially across firms and products see Levy et al. [1997] - consistent with observed heterogeneity in the frequency of price adjustment as found in empirical studies such as Dhyne et al. [2006], Dickens et al. [2007] and Dixon and Le Bihan [2012].

Note that *efficient* price dispersion is a *relative* concept of efficiency. It compares the actual distribution of prices to a corresponding model with flexible prices, perfect information and profit maximization[7]. The price dispersion that is efficient from the firms' point of view Δ_t^* could be inefficient from a social planner point of view because there are other inefficiencies in the economy (e.g. imperfect competition or asymmetric information) that cause welfare to fall below its social optimum. In fact it could be constrained (second-best) efficient to have inefficient price dispersion in order to help mitigate other uncorrected externalities. A prominent attempt to demonstrate this point is the burgeoning optimal inflation rate literature, where it is common to augment benchmark New Keynesian models capable of generating (inefficient) price dispersion, such as the basic Calvo model, with additional frictions in order to derive non-zero optimal inflation targets. [8]

The most general statement that can be made is as follows.

Theorem 1. *Under aggregate nominal rigidity and aggregate nominal heterogeneity there exists* $\bar{\pi}$ *such that if* $\pi_t \neq \bar{\pi}$ *then* $\exists\, t' \leq t$ *such that* $\Delta_{t'} > 1$

[7]With suitable modification of the RBC skeleton profit maximization could be relaxed to firm objective maximization to allow for among other factors risk aversion or behavioral factors as considered by for example Jaimovich and Rebelo [2007] Choudhary and Levine [2010], provided it did not induce direct dependence between today's optimal price and past optimal prices conditional on other shocks and parameters in the model.

[8]The first friction studied was the existence of non-interest bearing money which brought the deflationary forces of the Friedman rule into place see Khan et al. [2003], Adão et al. [2003], Schmitt-Grohé and Uribe [2004] and Schmitt-Grohe et al. [2007]. Subsequently, further imperfections in the product, goods and labor markets have been considered Collard and Dellas [2005], Pontiggia [2012] and Ikeda [2015], along with a binding lower bound on nominal interest rates Billi et al. [2011], Coibion et al. [2012], Eggertsson and Giannoni [2013], and Eggertsson et al. [2015]. As this literature only considers price dispersion owing to trend inflation and ignores the additional dispersion created by *stochastic* shocks there will be an upward bias in reported optimal inflation.It is beyond the scope of this paper to quantify this omission.

Proof. The result is trivial if $\Delta_{t-1} > 1$ so assume $\Delta_{t-1} = 1$. By hypothesis that $\pi_t = \bar{\pi}$ it follows that $P_t \neq (1+\bar{\pi})P_{t-1}$. Now we know from lemma 2 that $\Delta_{t-1} = 1$ implies that $\forall\, p(i) \in \Omega_{t-1}\ p(i) = P_{t-1}$. Therefore $\pi_t \neq \bar{\pi}$ requires that $\exists\ p(j) \in \Omega_t$ such that $\exists\ p_(j) \neq p_i$. The assumption of non-trivial price rigidity ensures that there exists $p(k) \in \Omega_{t-1}, \Omega_t$ from the first part $p(k) = p(i)$, therefore from the second part for some $p(j) \in \Omega_t\ p(k) \neq p(j)$ so by lemma 2 $\Delta_t > 1$. □

Note that this general result cannot be tightened to link inflation to *contemporaneous* price dispersion because this would not encompass models with Taylor contracts. The reason is that with Taylor contracts the price required to remove price dispersion can differ from that required to stabilize prices (zero inflation). This is because under Taylor contracts inflation is determined by a comparison between current reset prices and those they are replacing whilst price dispersion is determined by the difference between current reset prices and past reset prices that have *not* been replaced. Intuitively, non-zero current inflation can cancel out past price dispersion. Under Calvo where old reset prices never disappear the zero price dispersion and zero inflation reset price coincide so the coexistence must be contemporaneous. Appendix A offers a simple numerical example to clarify these points.

3.1.2 Stochastic Extension

The analysis extends naturally to a stochastic environment- all expectations are taken with respect to the ensemble distribution. The case where $\Delta = 1$ can be ruled out as a probabilistic description of the economy on what are effectively trivial restrictions on the stochastic environment affecting the pricing decisions of firms with prices that are not perfectly flexible.

Proposition 6. *For an economy with aggregate nominal rigidity and aggregate nominal heterogeneity* $\mathbb{E}\Delta_t > 1$ *unless* $\exists\ \bar{\pi}_t$ *such that* $\mathbb{P}(\pi_t = \bar{\pi}_t) = 1$ *and unless the shock process has a point mass at* $\pi_t = \bar{\pi}_t$ *then* $\mathbb{P}\Delta_t > 1$

Proof. These are applications of Jensen's inequality and Chebyhsev's inequality from proposition 5. □

In a stochastic environment we know that as there can only be one rate of inflation that can implement the flexible price case $\Delta = 1$ therefore any stochasticity of inflation ensures that $\mathbb{E}\Delta > 1$ by Chebyshev's inequality. Whilst, small perturbations local to this inflation rate make the probability $\Delta = 1$ arbitrarily small so $\mathbb{E}\Delta > 1$ by the Jensen's inequality argument in proposition 5. Therefore the non-stochastic steady state is improbable unless the shock process has a point mass at $\bar{\pi}_t$. The economic message is that concerns about taking approximations or analyzing welfare as though $\Delta = 1$, are much more general than the particular equilibrium concept and approximation strategies I present later on in sections 4 and 6.

3.2 Persistence Properties

This section analyzes the persistence properties of price dispersion in two familiar models: the Calvo model of Chapter 2 and mainstay of this paper and the Taylor framework discussed briefly in Chapter 2. In general price dispersion persists even if the shock process generating it is not present in all time periods. The most general statement can be made in the context of the benchmark Calvo model set out in Chapter 2.

3.2.1 Price Dispersion with Calvo Pricing

Remark 3. *If inflation π_t is ever non-zero in the Calvo model price dispersion $\hat{\Delta}_t > 0$ will exist in all subsequent periods.*

The result follows simply from applying Lemma 2 and noting that the set of prices in the economy Ω_t includes every previous price. This is because the fraction of prices in the economy equal to a given reset prices p_t^* never falls to zero no matter how far into the future one moves since $\iota_T(p_t^*) = \alpha^{T-t}(1-\alpha) > 0 \ \forall \ T > t$. This result has powerful implications for the class of equilibrium that can exist in a model with Calvo pricing. In particular it implies the current concept of equilibrium used in the literature (the non-stochastic equilibrium) does not arise in a Calvo model if it has ever had price dispersion.

Definition 4. The behavior of the New Keynesian model from time t can be representedby the continuation path $\mathcal{Z}_t^C = \langle Z_t, Z_{t+1}, \cdots \rangle$ where $Z_t = (\pi_t, y_t^e, \Delta_t, \pi_{t-1})$ which are governed by all the conditions set out in Chapter 2 apart from the policy rule equation (2.43).

The specific policy rule equation is omitted to allow consideration of the constraints on policy rules in general. Here policy is represented implicitly by continuation paths $\{\pi_t^C,\, y_t^C\}$, policy rules comparable to (2.43) but possibly time dependent could then by derived using the Euler equation (2.10).[9]

Definition 5. A non-stochastic continuation path from time t denoted $\bar{\mathcal{Z}}_t^C = \langle \bar{Z}_t,\, \bar{Z}_{t+1},\, \ldots,\, \bar{Z}_{t+\tau},\, \ldots \rangle$ is where $\forall\, \tau \geq 0 \quad \Pr(Z_{t+\tau}) = \bar{Z}_{t+\tau} = 1$, i.e. there is no uncertainty about future variables.

A *non-stochastic continuation path* corresponds to a perfect foresight model with initial value Z_t. As Z_t is dependent upon the shock carrying parameters from section 2 $\mathcal{U} = (\psi_T, \varphi_T, A_T)$ in a continuous fashion. The perfect foresight applies to the continuation of the shock processes also[10] $\times_t^C = \bar{\times}_t^C = \langle \bar{\Theta}_t,\, \ldots, \bar{\Theta}_{t+1},\, \ldots,\, \bar{\Theta}_{t+\tau},\, \ldots \rangle$ so $\Pr(\Theta_{t+\tau}) = \bar{\Theta}_{t+\tau} = 1, \forall\, \tau \geq 0$. A stronger concept is that of a *stable non-stochastic continuation path* defined as follows.

Definition 6. A stable non-stochastic continuation path denoted $(\bar{\mathcal{Z}}^*)_t^C$ is a non-stochastic continuation path where the shock process are held constant $\Pr(\Theta_{t+\tau}) = \bar{\Theta} = 1, \forall\, \tau \geq 0$

This corresponds to a non-stochastic model with initial position Z_t. Finally the strongest concept is that of non-stochastic equilibrium path from t.

Definition 7. A non-stochastic equilibrium path from t of a New Keynesian model denoted $(\bar{\mathcal{Z}}^{**})_t^C$ is a non-stochastic continuation path where $Z_{t+\tau} = \bar{Z}\ \forall\, \tau \geq 0$ i.e. all the variables remain constant in all future periods for sure.

In other words a non-stochastic equilibrium is a fixed point of the system where every future variable is certain to be constant at its present period value forever. This is the basic solution concept in microeconomics and growth theory. It is natural that macroeconomists wish to apply it to the New Keynesian model also. However,

[9] For convenience I use raw output y_t rather than the efficient output gap y_t^e introduced in Chapter 2 to characterize policy. The two formulations are equivalent as however the output gap variable is defined there must be a one-to-one mapping between them in a non-stochastic world.

[10] Formally there is a non-stochastic continuation path for Z_t if and only if there is a non-stochastic equilibrium for Θ_t with the *only if* following from the continuous dependence and the *if* following from the fact that Θ_t is the only source of uncertainty in the model.

whenever price dispersion is possible this is not in general correct even in the extreme case of a non-stochastic continuation path[11].

Definition 8. $\mathcal{Z}_t^H = \langle \cdots, Z_{t-1}, Z_t \rangle$ denotes the history of the variable Z up to time t

Let $\Delta(\bar{\pi}) = \Delta(\mathcal{Z}_t^H = \langle \cdots, \bar{\pi}, \bar{\pi} \rangle)$

Proposition 7. *In a model with Calvo pricing, a non-stochastic equilibrium* $(\bar{\mathcal{Z}}^{**})_t^C$ *with* $pi_T = \bar{\pi}$ *from time* t *will only exist if* $\Delta_t = \Delta(\bar{\pi})$*.*

Proof. The argument proceeds by contradiction suppose the converse then (21) takes the form of the following deterministic difference equation.

$$\Delta_T = \Delta(\bar{\pi}) + \vartheta^{T-t}(\Delta_T - \Delta_t) \tag{3.3}$$

Where

$$\vartheta = \alpha(1+\bar{\pi})^{\theta}$$

$$\Delta(\bar{\pi}) = (1+\bar{\pi})^{\theta} \frac{(1+\bar{\pi}-\alpha)^{\frac{\theta}{\theta-1}}}{(1-\alpha)^{\frac{1}{\theta-1}}(1-\alpha(1+\bar{\pi})^{\theta}}$$

$\Delta(\bar{\pi})$ is the non-stochastic steady state price dispersion and $\vartheta < 1$. This restricion is a requirement to ensure that a steady-state exists if not Δ would grow without bounds which would cause consumption C to tend to zero violating the transversality condition, equation (2.9)[12]. Now it is clear that if $\Delta_t \neq \Delta(\bar{\pi})$ it means Δ_T is time dependent contradicting the definition of a non-stochastic equilibrium for Z from time t. Note also that as $0 < \vartheta < 1$ the economy is converging monotonically towards its non-stochastic steady state but does not reach it in finite time. Therefore we know that $1 < \Delta_T < \max\{\Delta_t, \Delta(\bar{\pi})\}, \forall T > t$ □

For the zero-inflation steady state the behavior of Δ simplifies considerably to

$$\Delta_t = 1 - \alpha + \alpha\Delta_{t-1}$$

[11] Formally, a non-stochastic continuation path is necessary but not sufficient for a non-stochastic equilibrium. The necessity follows from noting that otherwise expectation errors could lead actual and expected values to diverge so that $\Pr(Z_{t+\tau}) = \bar{Z} < 1$ The non-sufficiency follows from the following counter-example.

[12] The threshold inflation rate $\bar{\pi} = (1/\alpha)^{1/\theta} - 1$. This restriction is well-known in the trend inflation literature. For developed countries it tends to be met quite easily see Ascari and Sbordone [2014]. More discussion on maximum steady state rates of inflation will be undertaken in Chapter 5 when stochastic equilibrium is introduced.

$$\Delta_t = \alpha^{t-t_0}\Delta_{t_0} > 1 = \bar{\Delta}$$

In other words the persistent behavior of the backward-looking price dispersion term stops the forward-looking variables output gap and inflation reaching equilibrium. Note however that the effect of initial price dispersion decays away when inflation is kept constant since $\lim_{T\to\infty} \Delta_T = \Delta(\bar{\pi}),\ \forall\, \Delta_t$. The (non-linear) Phillips Curve relation corresponding from 2.4 ensures y_t also has a limit (by the open mapping theorem once again) thus with a stable inflation policy the economy is ergodic $\lim_{t\to\infty} Z_T = \bar{Z},\ \forall\, Z_t$. The property that in the infinite limit a dynamical system "forgets" its initial position is called ergodicity. In section 4, the stochastic analogue of this concept will be used to define a meaningful notion of dynamic stochastic general equilibrium (DSGE). Take $t > t_0$. Since $\pi_t = \bar{\pi}$, $y_t = \bar{y}$ is not in fact an equilibrium of the system because it does not conform with all the optimization and market-clearing conditions that define the dynamical system laid out in section 2, to see this note that price construction equation (2.31) implies at each π_t maps to only one reset price p_t^*/P_t we can see this because the relationship between the optimal reset price and inflation is strictly monotonic- as

$$\frac{d\pi_t}{d(p_t^*/P_t)} = \frac{1-\alpha}{\alpha}\frac{(p_t^*/P_t)^{-\theta}}{(1+\pi_t)^{\theta-2}} > 0$$

Therefore the reset price will be constant in every period $p_t^*/P_t = p^*/P$. By recursively solving the optimal reset price equation for example (2.30). This implies real marginal costs $\varphi_t = \bar{\varphi}$ will also be constant. From the marginal cost expression equation (2.26) with no technology shocks $A_t = \bar{A}$ the real wage must be constant $W_t = \bar{W}$. Now note that in equilibrium the market-clearing condition implies when Δ decreases, labor l will decrease one-for-one which sets up a contradiction when we consider the optimal labor supply condition equation (2.11) by assumption the left hand side is constant but the right hand side must be decreasing- which completes the proof. Therefore persistence in the backward looking variable Δ will transmit to the other variables in the model. This mechanism is central to the analysis of the paper. Also note the following:

Corollary 3. *With Calvo pricing, there will always exist a non-trivial trade-off between inflation and output gap stabilization if there has ever been inflation variability.*

I have just shown that $\pi_t = \bar{\pi} \leftrightarrow y_t^e \neq \bar{y^e}, \forall t > t_0$ if there was initial price dispersion $\Delta_{t_0} > 1$ this implies that $y_t^e = \bar{y^e}, \forall t > t_0$ only if $\pi_t \neq \bar{\pi}, \forall t > t_0$. The link with inflation variability is provided by Lemma 2 it does not matter when the price inflation variability took place because of the permanence of price dispersion result Remark 1.

This is a profound result: it demonstrates that in a Calvo model there exists a non-trivial trade-off between inflation and output stabilization in any non-degenerate stochastic environment.

3.2.2 Price Dispersion with Taylor Contracting

This sub-section extends the results from the Calvo model to the Taylor contracting framework. Several results change but the theme that price dispersion generates staggered adjustment of the economy to shocks is retained. Furthermore, assuming the economy jumps to non-stochastic equilibrium immediately has misleading implications for the behavior of price dispersion with implications for welfare.

There is no analogue of remark 1 with Taylor pricing, as there is a maximum contract length so all contracts will eventually disappear from the price level- creating the possibility for price dispersion to disappear i.e. $\hat{\Delta} = 0$ if all the reset prices are the same for sufficiently long. Therefore Proposition 3 and corollary 2 apply only for as long as there exists prices set before $T = t_0$ that have not been reset. I focus here on the case of simple Taylor, Appendix A covers the simple extension to Generalized Taylor economy of Dixon and Le Bihan [2012]. Consider a simple Taylor economy where contracts last for M periods. There is staggered price adjustment so a fraction of firms $1/M$ are allowed to reset their price each period. In the knowledge that this price will remain fixed for exactly M periods. Therefore the price level construction equation takes the form:

$$P_t^{1-\theta} = \frac{1}{M} p_t^{*1-\theta} + \frac{1}{M} {p_{t-1}^*}^{1-\theta} + \cdots + \frac{1}{M} {p_{t-(M-1)}^*}^{1-\theta} \tag{3.4}$$

Firms set their reset prices as a weighted average of real marginal costs over the course of the contract so:

$$\max_{p_t(i)} E_t \sum_{T=t}^{t+M} Q_{t,T} \left[\frac{p_t(i)}{P_T} y_T - \varphi_T y_T \right] \tag{3.5}$$

Changes in the price level reflects the difference between the current reset price p_t^* and the price it replaced p_{t-M}^*.

$$P_t^{1-\theta} - P_{t-1}^{1-\theta} = \frac{1}{M}(p_t^*)^{1-\theta} - \frac{1}{M}(p_{t-M}^*)^{1-\theta} \tag{3.6}$$

from which can be derived the following expression for inflation

$$(1+\pi_t)^{\theta-1} = 1 + \frac{1}{M}\left(\frac{p_{t-M}^*}{P_t}\right)^{1-\theta} - \frac{1}{M}\left(\frac{p_t^*}{P_t}\right)^{1-\theta} \tag{3.7}$$

Now the evolution equation for Δ analogous to equation (19) in the Calvo model is

$$\Delta_t = \frac{1}{M}(p_t^*)^{-\theta} - (p_{t-M}^*)^{-\theta} + (1+\pi_t)^{\theta}\Delta_{t-1} \tag{3.8}$$

To see that a trade-off between inflation and output stabilization exists for the first $M-1$ periods I proceed by contradiction. Assume an equilibrium $Z_t = \bar{Z}$ exists from equation (3.6) you can see that to have a constant level of inflation $\pi_t = \bar{\pi}$ there must be a one-for-one relationship between p_t^* and p_{t-M}^*. Given $\Delta_{t_0} > 1$ there must be at least one price in the period t price level p_{t-j}^* where $0 < j < M$ such that $p_{t-j}^* \neq p_t^*$. When this price comes to be replaced at $p_{t-j+M}^* \neq p_t^*$ the optimal reset price equation (3.5) requires that $W_t \neq W_{t-j+M}$. Now by hypothesis $y_t = \bar{y}$ so from equation (2.26) the real wage and labor supply must move in the same direction- however this means aggregate income has increased- which violates equation (2.14)- the condition that all income must be consumed. From period $t+M$ onwards the economy reaches non-stochastic equilibrium as $\Delta = 1$. Therefore it takes precisely M periods for the Taylor economy to transition to non-stochastic equilibrium. This equilibrium is efficient i.e. equal to the flexible price output.The non-stochastic system is therefore ergodic. The result extends easily to the Generalized Taylor set-up where the non-stochastic equilibrium is reached after J periods- where J is the length of the longest contract.

3.2.3 Applications

This subsection considers a family of examples from macroeconomics that illustrate how ignoring price dispersion by using the ZINSS approximation to staggered pricing can generate spurious dynamics in a transparent non-stochastic environment. It is common to represent a New Keynesian economy with staggered nominal adjustment as switching between alternate non-stochastic steady states with no dynamic adjustment. This sub-section has proven that this approach is erroneous. The two models to which this approach has been most commonly employed have featured binding zero lower bounds on monetary policy and switches in monetary policy regime. It has also been used to model major exchange rate devaluation and structural adjustment episodes Uribe and Schmitt-Grohe [2012], Schmitt-Grohé and Uribe [2012a], Farhi and Werning [2012], Farhi and Werning [2014], Na et al. [2014], Eggertsson [2012] and Eggertsson et al. [2014].

Zero lower bound models seek to operationalize Keynes' idea of a liquidity trap. The idea is that arbitrage with money, which yields a zero nominal interest rate, prevents the Central Bank from cutting nominal interest rates below zero. Therefore if there is a sufficiently large fall in aggregate demand such that the desired nominal interest rate falls below zero, this zero bound will bind such that the economy will be demand-constrained with inefficiently low output.

Crucially, in a New Keynesian model because the representative consumer is forward-looking, liquidity traps cannot be permanent or the consumption problem explodes and the transversality condition is violated[13]. The literature on zero lower bound models is voluminous. They have been used to explain large economic contractions following financial crises and study optimal monetary and fiscal policy responses that might mitigate or overcome the constraint of the zero bound on nominal interest rates Corsetti et al. [2010], Lorenzoni and Guerrieri [2011], Eggertsson and Krugman [2012], Farhi and Werning [2013], Corsetti et al. [2013] and Correia and Farhi

[13] In an overlapping generation model with borrowers and savers this result does not apply as demonstrated by Eggertsson and Mehrotra [2014] although a permanent liquidity trap would appear implausible.

[2013][14]. Unsurprisingly, such models have become extremely popular following the global financial crisis of 2008 and the subsequent spell of near zero short interest rates across major industrialized economies.

The timing convention in these two steady state models is as follows: the economy starts at non-stochastic equilibrium- then there is an unanticipated shock. The shock has to be unanticipated or inflation would fall as the time of the zero bound spell approached in anticipation of the future deflation. To clarify this behavior would not contradict proposition 1- which is derived under the assumption that there is no zero bound on nominal interest rates. Finally leaning on the forward-lookingness derived in proposition 1 the model is closed with the economy jumping back to its non-stochastic steady state. This amounts to assuming the liquidity trap is a one-off (or occurs with vanishingly small probability) but would be valid if the probability of the economy transitioning from the normal-times benchmark equilibrium to a liquidity trap were sufficiently small to make the expected deflation associated with future zero bound spells of an order of magnitude less than or equal to the squared term in the series expansion of inflation[15]. This is incorrect. Many of the above models employ Calvo pricing which means that even if the economy begins at a non-stochastic equilibrium- it will never reach another one either whilst the zero bound is binding or afterwards when the shock has been turned off.

An alternative strategy has been to fix the length of the zero bound spell. The timing convention here is that the economy starts from non-stochastic equilibrium then experiences a shock known in advance to last T periods. The model is again closed with the economy jumping straight back to non-stochastic equilibrium (absent policy

[14] See also Woodford [2011b], Eggertsson and Giannoni [2013], Benigno et al. [2014], Denes et al. [2013], Eggertsson and Woodford [2004], Eggertsson [2006], Eggertsson [2011], Eggertsson et al. [2009], Adam and Billi [2007], Werning [2011], Cook and Devereux [2011b], Cook and Devereux [2011a], Cook and Devereux [2013], Araújo et al. [2013] and Schmitt-Grohé and Uribe [2014].

[15] Whether this alternative assumption is valid is difficult to gauge for two reasons. First it may be difficult to delineate zero bound spells- as their occurrence may be sensitive to changes in the policy environment- for example Eggertsson [2008] assumes the United States was characterized by a deflationary liquidity trap during the early phase of the Great Depression 1929-1933 on the grounds that under a stabilizing policy regime it would have been. However, in reality short term interest rates were significantly above zero throughout this time- which to me indicates the United States was not in a liquidity trap at this time. Similarly, World War 2 mobilization and financial repression measures designed to ease war financing makes it very difficult to ascertain the model consistent definition of the end of the zero bound spell- which is the point when private sector demand had fully recovered from the Great Depression shock see Reinhart [2012]. Finally, it has proved challenging to explain the length of zero bound spells, simultaneous with the small fall inflation- which calls into the question the validity of any parameter estimates.

changes) when the shock is turned off. There is no steady state whilst the zero bound binds, as resetters at different times face different length of zero bound spell relative to non-stochastic steady state equilibrium,so deflation will in fact moderate over the course of the zero bound spell. This approach is now the more popular Cogan et al. [2010], Christiano et al. [2011], Amano and Shukayev [2012], Erceg and Lindé [2014], Gertler and Karadi [2011] and Gertlera and Karadib [2013].

However, the approach in these papers still falls foul of the results in the previous sub-section. The dynamics during the zero bound spells will be wrong because many of these papers ignore the effect of the price dispersion associated with the deflationary shock implied by Lemma 2 and Corollary 1. Secondly, the economy will never return to steady state because of the assumption of Calvo pricing in each of these papers and Corollary 2.

Nevertheless, the major results from these papers would stand up as they do not depend on a particular specification of inflation dynamics. For example, Rognlie et al. [2014] and Korinek and Simsek [2014] show that liquidity traps generate deep recessions followed by recoveries even when the extreme old Keynesian assumption of no price changing during the liquidity trap is made. The concern is the dynamics the previous formulations highlighted.

The final two sections of this chapter are conceptual as opposed to practical the first links the restriction $\Delta \geq 1$ to non-negativity constraints and the presence of poles. The second locates the market failure associated with price dispersion as a coordination failure in the planning of production.

3.3 Non-Negativity and Poles

Non-negativity constraints are an important topic in dynamic economics. The ZLB model discussed in the last section are prominent in business cycle macroeconomics. Likewise (not necessarily zero) borrowing constraints are central to modeling financial markets and the distribution of income, two of the most popular topics in our discipline. Seminal contributions include Bewley [1977], Huggett et al. [1993], Aiyagari [1994] and Krusell and Smith Jr [1998]. Recent developments and applications include Moll [2014], Zambrano [2015], Kim and Khan [2015], Khan et al. [2016] and

Kaplan et al. [2018] as well as several ZLB papers mentioned previously. Irreversible investment presents similar challenges and a wide variety of applications see for example Arrow et al. [1970], Kogan [2001], Stern et al. [2008], Khan et al. [2013], Belo et al. [2014], Rozenberg et al. [2014] and Imura et al. [2015].

To derive a connection between these clear cases of non-negativity constraints and our restrictions upon price dispersion, we need to relax the notion of non-negativity to allow a model to be arbitrarily well approximated by a model with a non-negativity constraint.

Definition 9. A structural model $\mathbb{E}Z_{t+1} = f(\gamma, Z_t, U_t^S)$ has a weak non-negativity constraint with respect to a subset of the parameters $\gamma' \subset \gamma$ and variable $X = \tilde{f}(\gamma, Z_t, U_t^S)$ if for every open neighborhood in the relative topology of γ' there exists $\gamma'(0)$ such that there exists an $X^0 < 0$ where $Z_i(X^0, \gamma'(0)) < \inf \mathcal{Z}_i$ but $Z_{i'} \in \mathcal{Z}_i$ for all $X \in \mathbb{R}^k$ for some $Z = (Z_i, Z_{i'})$.

Theorem 2. *The non-stochastic version of the Calvo model with functional forms in place has a weak non-negativity with respect to* $(\theta, \eta, p_t^*, \Delta)$

Proof. First I create the parameter conditions for a binding non-negativity constraint then I show that these conditions are fulfilled arbitrarily close to any γ consistent with weak non-negativity. Note that if $\theta = a/(2n+1)$ for positive coprime integers $(a, 2n+1)$ then we know that Δ will be an odd function therefore since $\lim_{p^* \to 0} \Delta = \infty$ if we allowed $p^* < 0$ then we would have $\tilde{f} : p^* \to \Delta$ would be $\mathbb{R}/\{0\} \to \mathbb{R}$. However, the price level and therefore inflation will remain an even function so its state space will be unaffected by allowing negative reset prices.

Now suppose $\eta = b/2n$ for positive coprime integers $(b, 2n)$ then the marginal cost function will be even hence the state space of marginal costs will be unaffected by allowing negative prices. Therefore solving out the Phillips curve for sequences of output proves this must be true for output Y also. This covers all the variables Z. Therefore, we are complete if we can show that there are appropriate (θ, η) in every open neighborhood. In every open neighborhood it must be the case that $\theta^0 - \epsilon < \theta < \theta^0 + \epsilon$ for some $\eta > 0$. Since the rationals are dense in the reals we know that there must exist either appropriate $a/(2n+1)$ or $b/2n$ if there existed $a/(2n+1)$ then our conjectures about θ would be complete. However, this would not be the case if the interval only contained rationals of the form $b/2n$. Fortunately, this would generate a contradiction. Suppose that there was an open set containing only integers of the form $b/2n$ then there must exist $\epsilon > 0$ such that

$$|b/2n - a/(2n+1)| > \epsilon$$

for all (a, b, n) such that $(a, 2n+1)$ and $(b, 2n)$ were coprime but then we know that if consider n a continuous variable it follows that

$$\lim_{n\to\infty} |b/2n - a/(2n+1)| \to 0$$

This means that there must be some $\bar{n}(\epsilon)$ such that $n > \bar{n}(\epsilon)$ implies

$$|b/2n - a/(2n+1)| < \epsilon$$

now selecting the integer in for example $[\bar{n}(\epsilon)+1, \bar{n}(\epsilon)+2]$ completes the contradiction. □

Economically allowing negative prices makes no sense. They would contradict the underlying optimization problems. If we allowed stochasticity the expectations would not exist because the expectation integral would not be σ finite owing to its behavior local to the pole $p^* = 0$. However, I feel the result still presents an interesting point of comparison with existing models with clear non-negativity constraints.

3.4 Market Failure New Keynesian Model

This section addresses the source and consequences of market failure in the New Keynesian model. I show that price dispersion induces an externality between the marginal costs paid by the firm and the marginal cost of output to society. I show how this discrepancy leads to inefficiency in the aggregate economy. These results overturn the 'Divine Coincidence' associated with Blanchard and Galí [2007] and formulate the arguments of Damjanovic and Nolan [2010] Alves [2014] from a traditional microeconomic externality perspective.

These results give justification for the traditional understanding of Keynesian economics as the economics of *coordination failure* present in both early New Keynesian models such as Blanchard and Kiyotaki [1987], Cooper et al. [1988], Ball and Romer [1990] and Ball et al. [1991] and more traditional Keynesian work such as Leijonhufvud [1968], Leijonhufvud [1981]. It should draw attention away from Keynesian models with heterogeneous agents such as Kaplan et al. [2018] towards those with heterogeneity on the firm side.[16]

Although, my result is closer to the latter, as unlike the early New Keynesian studies there is no strategic interaction between individual firms as each is small relative to the market it takes seriously the idea that incomplete optimization rather than some aspect of the optimization environment is what causes market failure. The arguments extend the Acemoglu [2008] notion of a *representative firm* which he shows to be necessity for efficiency in a competitive economy. The following is an adaption of the original definition to allow for the possibility of non-price taking behavior in the output markets and the possibility for an uncountable set of firms, as we have here. The two preliminary results (theorem 3 and theorem 4) are of independent interest, as they do not rely on the existence of a specific lump sum tax- proportional subsidy scheme designed to remove the quantity distortions present under imperfect competition. They are therefore robust to second best concerns that one market failure monopoly markups might cancel out another nominal coordination failure. The first result demonstrates the proximate source of the market failure.

Theorem 3. *Whenever price dispersion $\Delta > 1$ there is an externality because the marginal social cost MC^s of producing a unit of output Y exceeds the private marginal cost mc^p.*

Proof. The marginal social cost can be calculated from the following problem

$$\min_{L} \ \mathcal{Z} = WL - \lambda(\Delta f(L) - \frac{\bar{Y}}{\Delta})$$

which minimizes the aggregate cost of production subject to the aggregate production constraint expressed in terms of target output. It is analogous to the firms

[16] To this end it is worth noting that their so called HANK model fails to generate endogenous persistence its dynamics appearing to mimic those of the conventional Calvo framework in Section 2.4- see Debortoli and Galí [2017], .

minimization problem except that aggregate replaces the individual production constraint.Taking the first order condition and apply the envelope theorem reveals that $MC^s = \lambda = \Delta W/f'(L)$ combining the expression for marginal private costs derived in Section 2.2.2 with the familiar result that with a competitive labor market $W = f'(l) = f'(L)$ completes the proof

$$MC^s = \frac{\Delta W}{f'(L)} = \Delta > MC^p = \frac{W}{f'(l)} = 1$$

. □

The next results clarifies this market failure is detrimental to social welfare.

Theorem 4. *When $\Delta > 1$ the economy is Pareto inefficient.*

Proof. Suppose that an allocation associated with $\Delta_0 > 1$ were Pareto efficient and proceed by contradiction. Now consider the repricing that keeps the price level P and wage rate W fixed but has $\Delta_1 < \Delta_0$, then by convexity of technology this output combination is feasible, by the existence of Δ we know that aggregate profit increases $\Pi(\mathcal{P}(\cdot_\infty)(\mathbf{w})) > \Pi(\mathcal{P}(\cdot_\prime)(\mathbf{w}))$ □

Now the task is to link this inefficiency to the representative firm theorem 5.4 p191-192 Acemoglu [2008]. To do so requires a definition of a representative firm that encompasses Acemoglu's case of perfect competition and a New Keynesian one where each firm is a price setter.

Definition 10. A representative firm devises a production plan $\{y_t(i)\}$ for each firm in the economy that maximizes total profit $\int_i \pi_i \mathrm{d}i$ subject to

i The input price W

ii The price level P

iii Market clearing conditions

The difference between the classical price taking world and its Keynesian lies with the final two conditions. In the classical world the representative firm must condition not just on the aggregate price level but on the competitive equilibrium price of each individual firm. However, there is only one aggregate market clearing condition because varieties are perfect substitutes. By contrast in the Keynesian world the market clearing condition is that each firm must satisfy its demand curve, whilst each can set its own price. In both cases there is a labor market clearing condition. In fact the representative firm's problem can be viewed as choosing Δ to maximize joint profits with $Y(\Delta)$ determined by the labor market clearing condition.

Theorem 5. *Consider the economy in section 2 with the monetary policy rule and pricing optimization protocol left unspecified. Suppose further a welfare maximizing government with access to a production subsidies and lump sum taxes. The economy is Pareto efficient if and only if there exists a representative firm which occurs if and only* $\Delta = 1$.

Proof. The only if chain Pareto efficiency $\rightarrow$ representative firm $\rightarrow \Delta = 1$ follows from the definition and the contradiction argument in the previous theorem. $\Delta = 1$ $\rightarrow$ representative firm follows likewise. To prove the final implication that $\Delta = 1$ $\rightarrow$ Pareto efficiency note that by proposition 5- firms will all have the same price and therefore the same desired mark up m so that a common per unit subsidy $\tau = m/(1+m)$ funded by the appropriate lump sum tax implements the equilibrium that would arise if there were perfect competition and this is Pareto efficient according to the first fundamental theorem of welfare economics- see for example theorem 5.5 p197-198 Acemoglu [2008]. □

In the case of the benchmark Calvo model where $\Delta = 1$ is the familiar non-stochastic steady state the subsidy would be $1/(\theta - 1)$. There are practical difficulties with this solution. Tax systems designed to resemble true lump sum taxes have been widely criticized as regressive Smith et al. [1991]. This could be fixed by varying the tax according to fixed attributes correlated with productivity although this has proven unpopular Mankiw and Weinzierl [2010][17]. Furthermore, there are political economy concerns relating to the legal doctrine of *attainder* . Attainder is the idea that laws passed by elected representatives should be impersonal. It prevents the legislature in this case in their role as designers of the tax system, from circumventing the court system who are responsible for administering justice to individuals. A truly individual

[17] All these results should be read as specializations of the second fundamental theorem of welfare economics- Acemoglu [2008] theorem 5.7 p199-203 and associated appendix item theorem A 25 provides a comprehensive treatment of this result.

specific taxation system would surely be judged to violate this doctrine.[18] Instead, lump sum taxes are a convenient fiction to allow us to compare the Keynesian universe where there are two market failures with the Classical setting where there are none in the way that is common with other market failures like missing markets and public goods.

In any case asymmetric information would cause serious problems with the subsidy component. Implementing the optimal subsidy would require the government to know all information relevant to the distribution of prices at the same time as the firms, including each firm's history of price resets. This is in fact all the valuable information economy. A government with all the valuable information is a central planner. In this context, the efficiency and benevolence assumptions would seem highly doubtful. In reality the information relevant to correcting coordination failure would surely constitute private information of the firms. In this sense price rigidity is a deep market failure reflecting a reluctance on the part of firms to fully transmit private information through the price mechanism. Hayek [1945] and Stiglitz et al. [1996] provide extensive discussion of these underlying themes in information economics.

[18] Article I Section 9 Paragraph 3 of the United States Constitution reads "No Bill of Attainder or ex post facto Law will be passed." Kurland et al. [1987]. For a popular perspective see Rehnquist [2007] p166, an example of a modern judgement is as follows "The Bill of Attainder Clause was intended not as a narrow, technical (and therefore soon to be outmoded) prohibition, but rather as an implementation of the separation of powers, a general safeguard against legislative exercise of the judicial function or more simply - trial by legislature." U.S. v. Brown, 381 U.S. 437, 440 (1965). Lehmberg [1975] documents attainders (mis)use under English law. Jones, Ashby. "Would an AIG-Bonus Tax Pass Constitutional Muster? (A Tribe Calls 'Yes!')" Wall Street Journal, 18 March 2009 and Clarke, Connor. "No Bill of Attainder... Shall Be Passed". The Atlantic, 16 March 2009- contain discussions of attainder in the context of measures proposed in the aftermath of the recent financial crisis. Almost all successful democracies or stable legal regimes have strong provisions against ex post facto laws pertinent here.

Chapter 4

Stochastic Equilibrium

This chapter builds the concept of stochastic equilibrium central to the thesis which is then applied for the basic Calvo New Keynesian model. The structure of the chapter is as follows. The first section provides an overview of the requisite mathematical machinery under the heading ergodic theory. Along the way there is substantial discussion relevant to practitioner and subsequent theoretical developments.

The second section focuses on the methodological significance of a statistical rather than purely optimality based definition of the equilibrium. Its connection with heterodox economics and equilibrium refinement in microeconomics feature. The third section formally proves the existence of stochastic equilibrium that is non-generate (not equivalent to a non-stochastic equilibrium)- valid for the Calvo New Keynesian setup- introduced in Chapter 2.

The final section characterizes the functional equations of the stochastic steady state of the benchmark Calvo model. Here there is discussion of statistical challenges that would arise when estimating the model. There is also application of comparative statics at the stochastic steady state this admits comparison with its non-stochastic counterpart. This approach can be viewed as an extension of non-stochastic comparative statistics and in particular the distributive comparative statics agenda which analyses- the effect of a change in a parameter on the distribution of some heterogeneous agent economy Acemoglu and Jensen [2010], Acemoglua and Jensenb [2013], Acemoglu et al. [2015] and Jensen [2017][1]. Alternatively, it can be seen as a compliment to previous analysis of monotone comparative statics under uncertainty in partial equilibrium such as Simonovits et al. [1995], Athey et al. [2002] and Wagener [2006].

4.1 Ergodic Theory

Ergodic theory is the branch of mathematics that studies the long run behavior of dynamical systems. It will be used to build the equilibrium concept I will use throughout the rest of the paper. It is necessary to present several definitions integral to ergodic theory and dynamical systems.

Definition 11. A measure-preserving dynamical system is defined as a probability space with a measure-preserving transform.

$$(X, \mathbb{T}, \Sigma, \mu, T)$$

such that

(i) $X = \prod_{i=1}^{k} X_i$ is a vector space and $\mathbb{T}$ the time index

(ii) Σ_i is a σ-algebra over X_i and $\Sigma = \prod_{i=1}^{k} \Sigma_i$ is the σ-algebra of the product measure on X

(iii) X is measurable with respect to μ with $\mu_i : \Sigma_i \times X_{-i} \to [0, 1]$ so $\mu_i(\emptyset_i) = 0$ and $\mu_i(X_i) = 1$ and $\mu = \prod_{i=1}^{k} \mu_i$

(iv) $T : X \to X$ is a measurable transformation with cocycle property which preserves the measure μ, formally $\forall\, A \in \Sigma$, $\mu(T^{-1}(A)) = \mu(A)$

Consult Arnold [2013] for a definition of the cocycle property. Initially I will focus on the case where T is the operator that sends the current state of the economy to its expected state in future periods so $T^s X_{t+s} = E_t X_{t+s}$. I will show that a fixed point

[1] These papers have taken a different approach to mine tending to eschew differentiability and uniqueness of equilibrium in favor of order-theoretic arguments. Interesting applications have included an extension of a Carroll et al. [1996] result concerning the effect of changing inequality in an economy with borrowing constraints Aiyagari [1994]. Although, so far it has proved difficult to extend this framework to include aggregate uncertainty. The powerful uniqueness results detailed here in Section 4.1 should help.

of T is equivalent to an equilibrium of the dynamical system. It is helpful to work with the sequence of values associated to the operator T formally:

Definition 12. A **trajectory** of a random dynamical system $T : X \to X$ from an arbitrary initial position X_0 is given by the set $\mathcal{O}_T = \{X_0, T(X_0), T^2(X_0), \cdots\}$

The definition of a trajectory supposes T is a homeomorphism so that T^{-1} is an inverse function. The analagous mapping if T^{-1} is instead set valued will be called a *path*. The following attributes characterize ergodicity:

Definition 13. Let (X, Σ, μ) and $T : X \to X$ be a measure-preserving transformation. We say that T is ergodic with respect to μ (or alternatively that μ is **ergodic** with respect to T if one of the following equivalent statements is true

(i) for every $E \in \Sigma$ with $T^{-1}(E) = E$ either $\mu(E) = 0$ or $\mu(E) = 1$

(ii) for every $E \in \Sigma$ with $\mu(T^{-1}(E) \,\Delta\, E)$ we have $\mu(E) = 0$ or $\mu(E) = 1$ where Δ denotes the symmetric difference $A \,\Delta\, B = (A \cup B)/(A \cap B)$

(iii) for every $E \in \Sigma$ with positive measure we have $\mu\left(\bigcup_{s=1}^{\infty} T^{-s}(E)\right) = 1$

(iv) for every two A and B of positive measure, there exists an $s > 0$ such that $\mu(T^{-s}(A) \cap B) > 0$

(v) Every measurable function (formally $j \in L^1(\mu)$) with $j \circ T = j$ is *almost surely* (a.s) constant.

Intuitively ergodicity is the property that a system forgets its initial position and that it has well-defined long-run behavior. It is a property of the system as a whole or a subset that behaves as a component of the whole process. Conditions (ii)-(iv) state that if you move far enough back or forward in time any two positions of the system will occur in probability. Consult Aliprantis and Border [2007] theorem 20.7 for a proof of these equivalences that can be easily adapted to the vector integral case using Dunford and Schwartz [1958] lemma III.II.I p184-85.

Ergodic measures have a powerful uniqueness property. Although, a state space may have multiple ergodic sets- a given ergodic set can only possesses one ergodic measure- for proof consult Hairer [2006] theorem 5.7 p 40-42. Throughout this paper I take the ergodic set to be the entire state space consistent with the informal formulation in Section 2. Thus any ergodic measure will be unique. Moreover, part (iv) of the definition of ergodicity can be interpreted as a weak restriction on hysteresis that is intuitive in a business cycle context. Crucially, ergodicity allows me to formalize the idea of stochastic steady state the correct concept of equilibrium for a *Dynamic Stochastic General Equilibrium* model. The following definition describes the progress of a dynamical system from an initial position:

Definition 14. The stochastic equilibrium of $\mathbb{E}_t X_{t+1} = f(X_t \ \gamma)$ is an ergodic fixed point of the dual operator $P : \mathfrak{P} \to \mathfrak{P}$ where $\mathfrak{P}$ is the space of probability measures.

Since the focus here is vector-valued integrals $\mathfrak{P}$ will be a space of multi-valued probability distribution functions[23]. By application of Sklar's theorem, I can write $\mu = (p_1(X_1), \cdots, p_n(X_n), C(X))$ where C is the copula mapping from the marginals p_i to the joint distribution $\mathbf{p}$ of the whole model. Furthermore, this joint distribution will be unique μ a.e.. Uniqueness will extend to the marginals and copula also. Therefore P can be represented by the stochastic kernel $\mathbf{p}(X_{t+1}) = \int \mathbf{p}(X_t, X_{t+1})\, \mathrm{d}X_t$ where $\mathbf{p}(X_t, X_{t+1})$ is the joint distribution of the current and next period state of the economy. Eventually I will search in the space of continuous distributions where the appendage μ a.e. can be removed and the elementary proof of Sklar's theorem from

[2] This restriction is worth discussing. The Lebesgue decomposition theorem states that every measurable self-map defined on a set of real numbers can be decomposed into discrete, continuous and singular continuous measures see Halmos [2013]. I focus on continuous measures because the natural setting for the New Keynesian model is a Euclidean manifold where convexity and calculus arguments can be employed. I overlook singular continuous measures which are associated with chaotic dynamics see Lasota and Mackey [1998] for a textbook exposition.
This is because if one refined the model to include a small amount of imperfect information then the dynamics of the model would change. Formally, I hypothesize there would be a discontinuity at full information if we augmented the model with a variable reflecting for example uncertainty about beliefs there would be a discontinuity at the full information limit. This seems economically unappealing- uncertainty is part of macroeconomic life and is the subject of a great deal of research. Chaotic dynamics occur when the measurable sets are badly mixed up with their compliments. With chaotic dynamics one loses the intuition that agents can be uncertain but still know the approximate state of the economy. This is most problematic in a policy context. Of course a measure also includes the non-stochastic case of a point mass. However, since each endogenous variable in the structural model has been assumed to be an absolutely continuous function of an independent and identically distributed random variable, the non-stochastic case will not be needed. In fact IID property is not needed it would be fine for example if the errors had a non-degenerate ergodic measure. However, the assumption that each endogenous variable be absolutely continuous with respect to the measure of some shock is indispensable. It is nonetheless heartening that this theory can incorporate the limiting case of a non-stochastic economy.

[3] I am using the Radon-Nikodym theorem- for proof consult Dunford and Schwartz [1958] theorem III.12.6 p 214-15.

Nelsen [2006] becomes valid.[4]

To analyze the stability properties of stochastic equilibrium it is necessary to introduce stronger restrictions known as mixing conditions:

Definition 15. For a measure-preserving transform $T : X \to X$ the following are defining conditions for every $A, B \in \Sigma$

i **Strong Mixing**

$$\lim_{s\to\infty} \mu(A \cap T^{-s}(B)) = \mu(A)\mu(B)$$

ii **Weak Mixing**

$$\lim_{s\to\infty} \frac{1}{s}\sum_{t=0}^{s} |\mu(A \cap T^{-s}(B)) - \mu(A)\mu(B)| \to 0$$

iii **Ergodic**

$$\lim_{s\to\infty} \frac{1}{s}\sum_{t=0}^{s} \mu(A \cap T^{-s}(B)) \to \mu(A)\mu(B)$$

Lemma 3. *Strong Mixing ⇒ Weak Mixing ⇒ Ergodicity*

Proof. Provided by theorem 20.11 and corollary 20.12 p 662-664 in Aliprantis and Border [2007]. □

This is the so called ergodic hierarchy which I will return to later. Note that whilst the convergence norm for the state space will be Euclidean for function space it be the norm formed from the standard statistical divergence metric $\frac{1}{2}|\mathbf{p}(X_t) - \mathbf{p}_0|$ for some $\mathbf{p}_0 \in \mathfrak{P}$.[5] There is a powerful connection between convergence to stochastic equilibrium and mixing. The following theorem indicates the power of ergodicity and binds it to existing notions of equilibrium.

Proposition 8. *A stochastic equilibrium μ is globally stable if and only if the operator T associated to the expectation operator in $\mathbb{E}_t X_{t+1} = f(X_t, \gamma)$ is **strong mixing**.*

Proof. The task is to show that for all $\mu_0 \in \mathfrak{P}$, $\lim_{s\to\infty} P^s(\mu_0) \to \mu$

'IF' part: Consider Definition 19 (i) set $B = A = X$ the state space. Note that

[4] For a selection of proofs applicable to the general case and accessible to an audience of mathematically trained economists consult Schweizer and Sklar [1974], Moore et al. [1975], Carley and Taylor [2002], Rüschendorf [2009] and Durante et al. [2013].

[5] For the divergence to constitute a valid metric $\mathfrak{P}$ must be restricted to continuous probability distribution functions or μ a.e. equivalence class.

by lemma 4 we know that T is ergodic by Definition 17 part (i) so $T^{-1}(X) = X \subseteq T^{-\infty} \subseteq X$ hence $T^{-\infty}(X) = X$. Note further that $\mu(B) = 1$ and denote $\mu(A) = \mu$ then we have that

$$\mu(T^{-s}X) \to \mu$$

hence for all $X_0 \in X$ the limiting measure is μ^* and therefore by the Reisz representation theorem we have $P^s(\mu_0) \to \mu = \mu^*$ by Definition 18 and the uniqueness of ergodic measures on a set the proof is complete.
'ONLY IF' part: Suppose the process were not strongly mixing then there exists sets $A, B \in \Sigma$ and set of points $X_0 \in X$ that belong to the trajectories that feed into B denoted $T^{-\infty}(B) \subseteq X$ such that

$$\mu(A \cap T^{-\infty}(B)) \nrightarrow \mu^*$$

hence by ergodicity of stochastic equilibrium there exists $X_0 \in X$ such that $\lim_{s\to\infty} \mu(A \cap T^s(X_0)) \nrightarrow \mu^*$ then by the Reisz representation theorem there exists a measure μ_0 corresponding to X_0 such that $\lim_{s\to\infty} P^s(\mu_0) \nrightarrow \mu^*$. □

With weak mixing this would extend μ a.e. on the space of distributions. Note that the stochastic processes $\{\mathbf{p}(X_t)\}$ and $\{X_t\}$ are weakly rather than strongly mixing- since we are using the expectation operator.Introducing seasonality (using raw data) would create the possibility for ergodicity without mixing at the expense of losing equilibrium interpretation of μ^* but I will save this for later work.
Note that the distinction between weak and strong mixing allows me to demonstrate generality of the stochastic equilibrium Calvo model over the conventional form isomorphic to the Rotemberg model by proposition 3. Consider the pair $(\{\mathbf{p}(X_t)\}, \{\mathbb{E}_t X_{t+1}\})$ which are isometrically isomorphic up to terms reflecting higher moments, this pair is not strongly mixing since there is no limiting value for $\mathbb{E}_{t+s}(X_{t+s+1})$. By Definition 17 (iv) it will reach μ a.e. point in the state and probability spaces with positive probability and infinitely often therefore this contrasts with the standard case where $\mathbb{E}_t X_{t+1} = \mathbf{0}$ and therefore $\mathfrak{P}$ is a singleton and the only self-map is the fixed point which trivially fulfills strong mixing. Previous work on dynamics around stochastic

equilibrium has tended to focus on the case where $\{\mathbb{E}_t X_{t+1}\}$ is strongly mixing associated with the macroeconomy converging towards an independent and identical distribution process. Therefore this research represents a significant technical advance[6]. The following is a powerful result because it allows me to analyze a random dynamical system as if it were in non-stochastic equilibrium.

Theorem 6. *(**Birkhoff**) For a measure-preserving transform $T : X \to X$ on (X, Σ, μ) and a function $j(X_t) \in L^1(\mu)$ with time average $\mathcal{A}_{\mathbb{T}}(j(X_0)) = \lim_{s\to\infty} \frac{1}{s} \sum_{t=0}^{s} j(T^s X_0)$ and ensemble average $\mathcal{A}_n = \int j \, \mathrm{d}\mu$ then the two coincide for μ a.e. X.*

Proofs for scalar measures can be found in Hairer [2006] theorem 5.2 p 38-40 and Walters [1975] p 32-36 applied component-wise. The economic interpretation is that in stochastic steady state all the higher moments and functions thereof that can be integrated with respect to the ergodic measure are in equilibrium. This is unlike the non-stochastic steady state where the variance, kurtosis etc. are set to zero. It encompasses all informal notions explored in the previous literature.

Remark 4. *Note that since $\boldsymbol{p}(X_t) \in L^1(\mu)$ we can view the ergodic distribution $\boldsymbol{p}(X)$ as a suitable non-linear average (GSM) of the sequence of probability distribution functions $\boldsymbol{p}(X_t)$*

It conforms with the GSM Definition 1 of Chapter 2. The first condition is met because μ is a valid probability measure and therefore integrates to one, the second condition is met by monotonicity, (uniform) continuity of the Lebesgue integral and the properties of a measure-preserving transform.

Recall that an ergodic measure need not exist. Recall the non-stochastic case discussed in section 3 where a sufficiently high rate of inflation can cause Δ to grow without bound. This result is intuitive, in New Keynesian economics we start off with a Classical framework in which the price mechanism acts to damp fluctuations and bring the economy to equilibrium subject to some regularity conditions on the shocks.[7] We restrict the operation of the price mechanism by forcing a fraction of

[6] See for example Buera et al. [2013],Moll [2014], Achdou et al. [2014],Gabaix et al. [2016] or previous work such as Lee et al. [1997].Note this should not be read as a criticism of these works they consider structural as opposed to business cycle factors where discrete shifts in steady state seem more plausible and also their models lack the steady state bifurcation of the main model here.

[7] The consumers problem is discussed extensively in Stokey [1989] whilst the firms problem in the flexible price framework is static so existence follows if the shock processes are sufficient for the existence of expected profit.

firms to keep the same price. It should be unsurprising that for certain parameter values associated with high and costly inflation, the system will explode. Note that if price dispersion grows without bound it follows from the production function and the fact labor supply is bounded by a time endowment that the transversality condition is breached and the model in section 2 has no solution.

Naturally, the practical significance of these cases are somewhat doubtful. State dependence in pricing decisions will restrict the growth of price dispersion as inflation increases. It is common for state-dependent pricing models such as rational inattention to have regularity conditions that ensure a well-behaved optimization problem regardless of the inflation process. I suspend the consideration of stochastic equilibrium in state-dependent pricing models for future work. Nevertheless, it is proper and instructive to study the conditions for existence of equilibrium. There are two approaches, one focuses on the stochastic process of the model and the existence of certain moments. The first uses Lyapunov functions a generalization of contraction mapping- see Stachurski [2003], Nishimura and Stachurski [2005] and Stachurski [2007].

Theorem 7. *Let $\mathbb{E}_{t+1}Z_{t+1} = f(\gamma, Z_t, \mathcal{U}_t^S)$ be a structural model of the form discussed in Section 2. If there exits a Lyapunov function $J : \mathcal{Z} \to \mathcal{Z}$ with compact level sets and constant $\rho \in (0,1)$ and vector of constants $\boldsymbol{C} >> 0$ such that $\mathbb{E}_t J(Z_{t+1}) = \rho J(Z_t) + \boldsymbol{C}$, $\forall (Z, \mathcal{U}^S)$ then a stochastic equilibrium exists.*

For proof apply theorem 4.21 and 4.22 to the product Lebesgue integral. Note that where Banach's fixed point theorem was applicable J as the identity map and any $\mathbf{C} >> 0$ would suffice.[8]

The problem is it would probably be difficult to construct Lyapunov functions for a New Keynesian model. Note that Stachurski [2007] (Proposition 3.1) cannot be applied to prove existence using simply log-linearization because the stability of the linearization will in general depend on parameters of the stochastic steady state whose existence one is purporting to prove. This is because the aim of New Keynesian macroeconomics is to generate non-monotonic impulse response functions (IRF) that

[8]The Reisz representation theorem was used to move between the expectation I have used here and the form employed by Hairer where Z_{t+1} is the independent variable.

match delayed response to monetary policy and other shocks observed in the data.[9] This prevents one applying Banach's fixed point theorem and it is often difficult to construct alternative Lyapunov functions in a multidimensional setting with a variety of IRF shapes. Hence I provide a different approach that relies solely on the asymptotic behavior of the IRF and an assumption innocuous in applied work. Furthermore, this condition should be verifiable and more intuitive for practitioners.

I will assume that $\mathbb{E}Z = \mathbf{0}$ otherwise it does not exist because for one or more element of Z the time average does not converge. This conforms with the common practice of describing business cycle models in deviation from a steady state and the use of filters to abstract from factors such as technical progress or level differences across units (firms sales, workers earnings, countries GDPs etc.) when undertaking business cycle analysis. There is no loss of generality as the two will be isomorphic in large sample. Conversely, just because its sample moment $\mathcal{A}_s(Z) = \sum_{s=0}^{S}$ is zero does not mean Z_t represents an equilibrium (or more generally deviations from an equilibrium) since this time average might in fact diverge $|\mathcal{A}_{\mathbb{T}}| = \infty$ hence I use the terminology *candidate* to describe the empirical complement of a stochastic steady state, since an empirical estimate of a stochastic steady state can be constructed even if none exists for the underlying dynamical system.

To define concepts related to the IRF it is necessary to partition endogenous and exogenous into state (predetermined) variables and jump variables which are non-predetermined $Z_t = (Z_t^s, Z_t^j)'$ and $U_t^{\mathcal{S}} = (U_t^s, U_t^j)'$. Recall that a variable is predetermined if its next period value is known today so $\mathbb{E}_t X_{t+1} = X_{t+1}$ in the benchmark model all predetermined variables will be lags. Take lagged inflation for example and

[9] There are two familiar examples. The first is the response of output and inflation to an unexplained monetary contraction. This reaches a peak after four to eight quarters before dyeing away and losing statistical significance after around twelve to sixteen quarters- see Romer and Romer [2004] figure 2 p 1071 and note similar results in Christiano et al. [1999], Bernanke et al. [2005], Uhlig et al. [2005], Croushore and Evans [2006] and Coibion [2012]- intuitively the increasing part of the hump that stops us applying contraction mapping. Secondly, it is possible for the initial response to have a counterintuitive sign that is later overturned. The two examples are two well known examples. Firstly, the price puzzle where inflation initially *rises* in response to an interest rate increase. Although the evidence is not uncontroversial, there are several potential economic mechanisms- see Thapar [2008], Krusec [2010], Castelnuovo and Surico [2010], Kaufmann and Lein [2012], Rusnák et al. [2013] and Demiralp et al. [2014] and Ali and Anwar [2016] for empirical perspectives, read Ravenna and Walsh [2006] for discussion of a possible theoretical mechanism. The second example of contractionary technology shocks has been discussed previously. In fact it is uncommon for a large empirical macroeconomic model such as Christiano et al. [2005], Smets and Wouters [2007] or those kept by a Central Bank to produce IRF suitable for contraction mapping. Giesen et al. [2012], Burgess et al. [2013] and Kamber et al. [2016] lay out typical Central Bank style models that are even larger.

note that $\mathbb{E}_t(\pi_{(t-1)+1}) = \mathbb{E}_t\pi_t = \pi_t$ since current inflation belongs to the present information set of the firms' and representative household.For the subsequent theorem I require the following

Definition 16. The family of **Candidate Impulse Response Functions** Φ consists of all possible trajectories $\mathcal{O}_T(\mathbf{0}, Z^j, \mathbf{0}, u^j)$.

A particular impulse candidate response function will be denoted (Z^j, u^j) and $T(Z^j, u^j)$ will refer to the associated trajectory of the structural model. A *candidate impulse response function* maps the response of the model to shocks given that it started out from its (unique up to isomorphism) *candidate steady state.*

Definition 17. A structural model possesses an **Asymptotically Uniformly Contractive Impulse Response Function** if there exists a contraction-modulus- lag length pair (ρ, S) such that for all $s \geq S$ and any candidate pair $(Z^0, u^0), (Z^1, u^1) \in \Phi$

$$d(T^{s+1}(Z^0, u^0), T^{s+1}(Z^1, u^1)) \leq \rho d(T^s((Z^0, u^0), T^s(Z^1, u^1))$$

where $\rho \in (0, 1)$ and d is the relevant Euclidean metric on $\mathcal{Z}$.

This condition is an extension of contraction mapping that accommodates hump-shaped impulse response functions. It also utilizes the *regenerative approach* to Markov chains[10], the idea that for a finite dimensional state space model the system will always 'forget' its current position a finite number of periods into the future. This generalizes from the Candidate Impulse Response Function trajectories to the whole state space by the accessibility assumption. The economic argument is that however big the shock after some time $t + s$ we know that the economy contracts towards equilibrium uniformly across its state space.

Theorem 8. *Structural models such as those laid out in Section 2 will posses a stochastic equilibrium if they have Asymptotically Uniformly Contractive Impulse Response Functions.*

Proof. We know that Z_t is k-dimensional therefore the maximum number of lags it could posses is $k - 1$ corresponding to the case where Z_t were a univariate $AR(k - 1)$

[10]Consult Meyn and Tweedie [2009] for a brief introduction p506-508 Ross [2010] p. 442- 621 for a longer exposition and extensive discussion of basic applications in the operations research literature. To my knowledge this is its first application in economics.

model. By the accessibility assumption. the state space of T^k would be equal to the full state space $\mathcal{Z}$. Note that as it forms a connected manifold the space of trajectories associated to the asymptotically contractive impulse response $\mathcal{Z}$ forms a Banach space. Therefore apply contraction mapping point to the subsequence of trajectories starting at time $t + S$ $T^S(.)$. This yields a unique fixed point for the operator T representing an invariant mapping for $\mathbb{E}Z$. To prove this fixed point is ergodic note that deviations from each component time average $\mathbb{E}Z_i$ can be bounded by the following inequality that uses the contraction and the properties of a Euclidean norm.

$$-\lim_{s\to\infty}\frac{1}{s}\frac{S}{1-\rho}\max_{0\leq s\leq S}\|T^s(Z^0,(u^j)^0)\| \leq \mathcal{A}_{\mathbb{T}}(Z_i) \leq \lim_{s\to\infty}\frac{1}{s}\frac{S}{1-\rho}\max_{0\leq s\leq S}\|T^s(Z^0,(u^j)^0)\|$$

By the squeeze (sandwich) theorem $\mathcal{A}_{\mathbb{T}}(Z_i) \to 0$ and therefore $\mathcal{A}_{\mathbb{T}}(Z) \to 0$. The convergence is pointwise hence T is strongly ergodic. Recall that the model in deviations from steady state corresponding to $\mathbb{Z} = 0$ is isometrically isomorphic to the levels model where $\mathbb{Z}_t > 0, \forall t$ which forms a positive linear operator. By the Reisz representation theorem see Aliprantis and Border [2007] p496 (14.12) there is a corresponding measure that is also ergodic and invariant. It therefore constitutes a stochastic equilibrium. □

Note that this condition is not *necessary*, it is possible to have a sub-geometric rate of convergence to the invariant measure Nummelin and Tuominen [1983], Tuominen and Tweedie [1994] or Douc et al. [2004]. In macroeconomics this would allow for thicker tails and stronger persistence consistent with narratives related to hysteresis following shocks or slow recoveries for example following financial crisis. The following result treats these cases. First I require a relaxation of the uniformity inherent in Definition 21.

Definition 18. A structural model possesses an **Asymptotically Contractive Impulse Response Functions** if there exists a lag length S such that for all $s \geq S$ and any candidate pair (Z^0, u^0), $(Z^1, u^1) \in \Phi$

$$d(T^{s+1}(Z^0,u^0), T^{s+1}(Z^1,u^1)) < d(T^s(Z^0,u^0), T^s(Z^1,u^1))$$

This allows for the possibility that the contraction rate becomes arbitrarily close to one. Series in which $\lim_{s\to\infty} p(T^s, T^{s+1}) \to 1$ can either converge or diverge.[11] However, an assumption of bounded impulse response functions allows us to rule out the latter.

Theorem 9. *If for every $(Z^0, u^0) \in \Phi$ there is an upper bound on the magnitude of the deviation from equilibrium $\bar{Z}(Z^0, u^0)$ then the assumptions in Section 2 are sufficient for the existence of a stochastic equilibrium that is globally stable.*

Proof. The proof consists of an slight adaption and verification of theorem 5.2 from Stachurski et al. [2002]. Note that the every trajectory $\mathcal{O}_T(Z^0, u^0)$ lies inside a compact set $[-\bar{Z}(Z^0, u^0), \bar{Z}(Z^0, u^0)]$ this meets Definition 5.1 for Lagrange stability. Contractive Impulse Response Function is consistent with Definition 5.2 in his paper by accessibility assumption we know that the trajectory space and state space are equivalent and therefore the impulse response function is defined on a topologically closed set and the theorem can be applied. □

The final sufficient condition is perhaps the most economically intuitive. The following definition is crucial.

Definition 19. A sequence of probability measures $\{\mu_s : s \in \mathbb{T}\}$ is called **tight** such that for every $\epsilon > 0$ there exists a compact set $C \subseteq X$

$$\lim_{s\to\infty} \inf \mu_s(C) \geq 1 - \epsilon$$

A dynamical system T will be called **bounded in probability on average** if for each initial condition $X_0 \in X$ the sequence of sample averages $\{\mathcal{A}_s(X_0) : s \in \mathbb{T}\}$ is tight.

The last part implies that the measure of the limiting average $\mathcal{A}_{\mathbb{T}}$ has compact support. This condition encompasses the *asymptotically bounded mean* criterion that forms Definition 3 of Szeidl [2012] and forms the backbone of Proposition 1 and Theorem 1 of that paper. Note this approach also encompasses results used in earlier seminal work such as the seminal Hopenhayn and Prescott [1992]. It will feature in the existence proof in the next section.

[11] For example $\sum_{s=1}^{\infty} 1/s^2 = \pi^2/6$ with ratio limit $\lim_{s\to\infty} p(T^s, T^{s+1}) = s^2/(s+1)^2 \to 1$ on the other hand $\sum_{n=1}^{\infty} 1 = \infty$

Theorem 10. *A model defined by a finite dimensional Harris recurrent Markov chain has a stochastic equilibrium under the assumption that Z is bounded in probability on average.*

Proof. The proof revolves around theorem 12.0.1 in Meyn and Tweedie [2009], part (i) establishes existence of the invariant measure and part (ii) global stability. The proofs of these results are completed by two further theorems 12.1.2 and 12.4.1 respectively these encompass p286-7 and p297-8. The condition for existence is that the chain be weak Feller the property that T sends continuous functions to continuous functions see p128. To show that the structural model is weak Feller take the canonical form introduced in Section 2 and convert into the random function form $Z_{t+1} = f'(Z_t, U_t^S)$ note that the left hand side f is (strongly) differentiable by the fundamental theorem of calculus since it is composed of Riemann integrals and hence the chain rule ensures the derivative of the left hand side exists since $\mathbb{E}dZ_{t+1}/dZ_t = \mathbb{E}dZ_{t+1}/dZ_{t+1} \cdot dZ_t/dZ_{t+1}$ where for each $1 \leq i \leq k$, $\mathbb{E}dZ_{i,t+1}/dZ_{i,t} = 1/(dZ_{i,t}/dZ_{i,t+1})$ and a similar argument with respect to U_t^S shows that f' is in fact differentiable and therefore continuous also. Hence, the original system T is weak Feller by Proposition 6.1.2 p130 and the proof of equilibrium existence goes through. □

The economic interpretation is powerful business cycle models are formulated in terms of gapped variables which are deviations from a trend or average that is interpreted as an equilibrium. For this interpretation to carry weight it must be the case that we can bound long-term averages otherwise the business cycle intuition of the model is lost. Furthermore, the Feller property is a weak restriction for a large economy. In macroeconomic terms these are regularity conditions except for the case of a random walk which will be treated next.

The challenge in empirical work is that prominent level variables in particular output often do not display ergodic behavior. There are two routes around this problem. The first is to filter the data. It is common econometric practice to analyze data that has been transformed to remove effects such as seasonality and technical progress that structural models commonly abstract away. [12] Let X_t be the filtered data

[12]There is no guarantee that the structural model that leaves out the non-stationary features estimated on filtered data will be isomorphic to that structural model augmented with the non-stationary features and confronted with non-filtered data. Indeed, there is evidence the two behave

from the structural model Z_t and $\mathcal{F}(t, \{Z_t\})$ the filter so that $X_t = \mathcal{F}(Z_t)$ which forms a homeomorphism. The most desirable connection is an isometric isomorphism between the operators $\mathcal{P}_X : \mathbf{p}(X) \to \mathbf{p}(X)$ the ensemble probability distribution over X the detrended model and $\mathcal{P}_{\mathcal{F}} : \mathbf{p}(Z) \to \mathbf{p}(Z)$ the ensemble probability distribution over the actual model. This is possible if $\mathcal{F}$ were deterministic or in large sample if it could be consistently estimated for example if $\mathcal{F}$ were a time trend with finite parametization or a suitable non-parametric procedure were employed. An example is in order consider the simplest univariate model with linear trend and an exogenous error

$$Z_t = a + bT + X_t + u_t$$

where u_t $N(0, \sigma_u^2)$ and X_t $N(\mu_X, \sigma_X^2)$ then we know by the basic properties of the normal distribution that Z_t $(a + bT + \mu_X, \sigma_u^2 + \sigma_X^2)$ hence $\mathcal{F} = Z_t - a - bT - u_t$ if the parameters a and b were known. The isometry would take the form $\mathbb{I} : \mu(\mathcal{P}_X) \to a + bT + \mu(\mathcal{P}_{\mathcal{F}})$ In large sample unknown a and b could be replaced by their OLS/ML sample analogs and consistency would follow from a basic law of large numbers argument subject to standard regularity conditions. It is common however, to use a stochastic trend to reflect factors such as shifts in policy regime or omitted structural change. This breaks the isometry because the mapping $\mathcal{F}(u^{\mathcal{T}})$ now depends on the unobserved trend disturbance $u^{\mathcal{T}}$. Likewise the ergodic distribution of X would depend on parameters $\gamma^{\mathcal{T}}$ of the distribution of $u^{\mathcal{T}}$ and cause identification problem for the parameters of the underlying structural model γ.

There are two strategies to circumvent this problem. The first is to specify a functional form for the trend usually either the linear as above or first difference so $\mathcal{F}(Z_t) = Z_{t-1}$. The other is to calibrate hyperparameters that control the stochastic trend such as the smoothing parameter λ in the Hodrik-Prescott filter[13] upper and lower frequencies in the bands or persistence and variance terms in drifting parameter models. These are commonly selected to accord with widely held intuitions about the business cycle associated with the dating methodology introduced by Burns and

differently see Olivei and Tenreyro [2007], Olivei and Tenreyro [2010], Tesfaselassie [2013] and Snower et al. [2017] however such trade-offs are present in all modeling and I leave these ideas for future research.

[13] In statistics the technique is named after Whittaker [1922].

Mitchell [1946] and the pronouncements of official bodies such as the National Bureau of Economic Research in the United States[14] Subject to these restrictions on the stochastic trend stochastic equilibrium becomes a testable restriction. There are a battery of tests parametric and non-parametric for ergodicity of certain moments- see for example Kearns et al. [1997], Ling and McAleer [2002], Fedotenkov [2013], Fedotenkov [2014] and Trapani [2016]. The difficulty lies with the fixed point property and its statistical counterpart the mixing condition. In this section I have argued it must hold for the theoretical model, for example the output and inflation gaps by the maintained assumptions on the shock process in Chapter 2.

However, whether it holds for any estimates thereof is an empirical question. The first difficulty is that filters tend to have poor small sample properties Jin et al. [2015] and de Jong et al. [2016] demonstrate that for a wide class of processes the most popular filter Hodrik Prescott often fails to eliminate stochastic trends in finite time. Therefore it may be advisable to test for stationarity in small samples.

The more serious problem is that the filtration can radically alter the dynamic properties of a system encoded in its stochastic equilibrium. Early work with the RBC model showed that the Hodrik Prescott filter could generate rich dynamics even where none exist in the underlying data- see Harvey et al. [1993], King and Rebelo [1993], Cogley and Nason [1995a] and recent derivations in Jin et al. [2015]. Filtration can

[14]The widely used Hodrik-Prescott filter decomposes a time series into a trend and deviation component according to the following objective function $\mathcal{L} = \sum_{s=-1}^{S}(Z_s - X_s)^2 + \lambda \sum_{s=-1}^{S}\{(X_s - X_{s-1}) - (X_{s-1} - X_{s-2})\}^2$ with the initial four values for X_s predetermined. λ controls the variability of the trend relative to the deviations from trend, with lower λ implying a more flexible trend. In particular $\lambda \to 0$ corresponds to $\mathcal{F}Z_t = Z_t$ the case where the trend absorbs all variation in the series whilst $\lambda \to \infty$ imposes a linear trend.

The practice in the literature has been to set $\lambda = 1600$ for quarterly frequency this corresponds to typical average cycle lengths of four- to six years see Hassler et al. [1992] and Machado [2001] for calculations and Ravn and Uhlig [2002] for adjustments to different data frequency. With the band-pass filter, the average cycle length is determined as follows. First, admissible frequency bounds are determined then the average within the bounds is determined by the data- see Baxter et al. [1999] who consider business cycle frequencies as between six and thirty two quarters. The two approaches are mutually reinforcing as smoothing parameter choices are motivated by the implications for cycle length, although this was not the original approach of Hodrick and Prescott [1997]. Kourtellos and Stengos [2010] touches on similar issues in the drifting coefficients literature.

These parameter choices are not uncontroversial however Hamilton [2017] argues that the value of 1600 is unrealistic because it is several orders of magnitude away from the maximum likelihood estimate of λ from a Gaussian state space model with white noise drift in the errors which he regards as a plausible specification for the underlying economic dynamics. Moreover, factors such as long memory in financial markets or financial and fiscal cycles that appear to be longer than typical post war business cycles may be distorted by cyclical decomposition that favor a shorter cycle with likely spillovers to the dynamic properties of macroeconomic models such as the one here that abstract from these real world features. Consult Claessens et al. [2011], Claessens et al. [2012], Landmann [2012], Fernández-Villaverde et al. [2013], Reinhart et al. [2014] and Diebold et al. [2013].

influence other time series properties of the data such as business cycle turning properties of the model or shapes of impulse responses- see Canova [1994] and Canova et al. [1998].

The take home message is that in applied work what is needed is to seek robustness across methods combined with willingness to take stance on the nature of any trends[15]. Overall stochastic equilibrium is a versatile and robust conceptual framework for empirical research but filtering is necessary to apply it to a macroeconomic environment with trends and the possibility for structural breaks[16].

4.2 Existence of Stochastic Equilibrium

Here I use Theorem 11 along with two boundary conditions to prove that a unique stochastic equilibrium will arise and will be non-degenerate (not equivalent to a non-stochastic equilibrium.) In return for mathematical rigor unlike the previous section the arguments here are specific to the Calvo framework.

Theorem 11. *The Calvo model described in subsection 2.4 and characterized by proposition 1 possesses a unique stochastic equilibrium μ^* provided that social welfare is not minimized i.e.* $\lim_{T\to\infty} U(C_T, l_T) = -\infty$

Proof. From theorem 11 the task set before us is to prove that $Z_t = (\pi_t, y_t, \Delta_t)$ is bounded in probability on average. To bound y note that the time constraint implies that $y_t \leq \bar{l} A_t$ where $\bar{l}$ is the maximum labor supply (the entire time endowment.) Now assumption 1 accessibility and existence parts allows the application of a strong law of large numbers to the shock process of A combining with the Inada condition implies $\mathbb{E} y_T$ is asymptotically constrained to the set relatively compact $(0, \bar{l}\mathbb{E}A]$. The strategy is a little different for π we do not need to worry about infinite time limits we can bound for all time as follows express π in terms of the relative reset price p_t^*/P_t

$$\pi_t = \left(\frac{1}{1-\alpha} \left(1 - \frac{\alpha}{(p_t^*/P_t)^{\theta-1}} \right) \right)^{1/(\theta-1)}$$

[15] Consult Burnside [1998] for an insightful and still pertinent discussion of issues with detrending. In fact it appears that the most robust method is to treat time series such as output as if they were random walk Christiano and Fitzgerald [2003].

[16] Structural breaks a popular description of one off policy changes are conceptually inconsistent with ergodicity. However, an alternative strategy allowing regime switching can be squared with ergodicity although none of the stability or convergence results derived here would be applicable. I leave this for future research.

Inflation is clearly strictly increasing in the relative reset price. Therefore by taking the infinite limit we have a strict upper bound on π. By definition $\pi \geq -100\%$ and this equality is strict with the lower bound a pole of the Euler. Therefore by the basic properties of the Lebesgue measure we know that $\mathbb{E}_t \pi_T$ for any T is constrained to the relatively compact set $(-1, 1/(1-\alpha)^{\theta-1} - 1)$. The proof will be complete if we can show that the final variable of the cocycle $\mathbb{E}\Delta$ is also constrained asymptotically to a compact set.

Assume the converse and proceed by contradiction. The expectation must explode to positive infinity $\lim_{T\to\infty} \mathbb{E}_t \Delta_T = \infty$. This follows from continuity of the cocycle and precompactness of the space (this means that every open set can be approached as the limit- it holds for all manifolds). Now consider resource constraint I have established that the right hand side is bounded by $\bar{l}\mathbb{E}A$ this implies $\lim_{T\to\infty} \mathbb{E}_t Y_T = 0$. Then the Inada condition and basic properties of the Lebsesgue measure imply that the instantaneous utility $\lim_{T\to\infty} \mathbb{E}U_T = \phi_T(u(C_T) - \nu(l)) = -\infty$ a contradiction. Non-degeneracy follows from the isomorphism of the cocycle to the non-degenerate shock processes. □

The restriction that price dispersion does not grow without bound $\mathbb{E}\Delta < \infty$ implies a novel restriction on the inflation process that will be explained in the next section.

4.3 Characterizing the Steady State

This section completes the characterization of the stochastic steady state by solving for the equations satisfied by Y and Δ. This allows me to solve for the steady state values of quantities such as profit Π and welfare U. Finally, I undertake comparative statics with focus on the existing non-stochastic steady state formulations.

Focus first on the optimal price setting equation and the weighted marginal cost $\aleph$. Birkhoff's theorem tells us that when an economy is in stochastic steady state every expected function of future variables is in equilibrium- at its long-term average value.

$$\aleph_t(\pi, Y, \Delta) = u'(Y)Y\frac{\nu'(\Delta Y/A)}{A} + \alpha\beta\mathbb{E}\sum_{T=t+1}^{\infty}(\alpha\beta(1+\pi)^{\theta})^{T-(t+1)}u'(Y)Y\frac{\nu'(\Delta Y/A)}{A} \tag{4.1}$$

where the expectation of the $\aleph_{t+1}$ expression is taken implicitly with respect to the ergodic measure. Now this can be expressed as an iterated integral with respect to the ergodic measure.[17]

$$\mathbb{E}\aleph_{t+1} = \sum_{T=t+1}^{\infty} (\alpha\beta)^{T-(t+1)} \int \cdots \int (1+\pi)^{\theta(T-(t+1))} u'(Y) Y \frac{\nu'(\Delta Y/A)}{A} \mathrm{d}\mu^{T-(t+1)} \tag{4.2}$$

Now with Tonelli's theorem Aliprantis and Border [2006] theorem 11.28 p419 and the fixed point property of the ergodic measure we can write the term from date $T-t$ term as

$$\left((\alpha\beta)\int (1+\pi)^{\theta}\mathrm{d}\mu\right)^{T-(t+1)} \left(\alpha\beta \int (1+\pi)^{\theta} u'(Y) Y \frac{\nu'(\Delta Y/A)}{A}\mathrm{d}\mu\right)$$

where the left hand term represents expected gross inflation cumulated from the present up to period $T-(t+1)$, whilst the right hand term is the contribution of marginal costs at time $T-t$ both evaluated at the ergodic measure. This forms a geometric progression with initial value the right hand term and ratio $\alpha\beta\mathbb{E}(1+\pi)^{\theta} < 1$ otherwise the sum diverges and the steady state does not exist. A tighter bound will appear later. Hence, the expression for $\aleph$ in stochastic steady state is

$$\aleph(Y,\pi,\Delta) = u'(Y) Y \frac{\nu'(\Delta Y/A)}{A} + \alpha\beta \frac{\mathbb{E}(1+\pi)^{\theta} u'(Y) Y \nu'(\Delta Y/A)/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}} \tag{4.3}$$

By the same method for the weighting equation

$$\daleth(Y,\pi,\Delta) = u'(Y)Y + \frac{\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}u'(Y)Y}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}} \tag{4.4}$$

Combining gives the implicit functional equation for output

$$\left(\frac{1-\alpha}{1-\alpha(1+\pi)^{\theta-1}}\right)^{1/(\theta-1)} =$$

$$\frac{\theta}{\theta-1}\left(u'(Y) Y \frac{\nu'(\Delta Y/A)}{A} + \alpha\beta \frac{\mathbb{E}(1+\pi)^{\theta} u'(Y) Y \nu'(\Delta Y/A)/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right) \tag{4.5}$$

[17] This is formally speaking a *weak integral* in particular a Bochner integral defined on the space of the Banach space of forward iterations of the cocycle f and the adjoint space formed by taking $\mathbb{E}_t$. For details consult Aliprantis and Border [2006] chapter 11.9 p428-431.

$$/\Big(u'(Y)Y+\frac{\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}u'(Y)Y}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\Big)$$

Viewed as a moment condition we have

$$\Big(\frac{1-\alpha}{1-\alpha(1+\mathbb{E}\pi)^{\theta-1}}\Big)^{1/(\theta-1)}=$$

$$\frac{\theta}{\theta-1}\Big(u'(\mathbb{E}(Y))\mathbb{E}Y\frac{\nu'(\mathbb{E}\Delta\mathbb{E}Y/\mathbb{E}A)}{\mathbb{E}A}+\alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}u'(Y)Y\nu'(\Delta Y/A)/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Big) \tag{4.6}$$
$$/\Big(u'(\mathbb{E}Y)\mathbb{E}Y+\frac{\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}u'(Y)Y}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\Big)$$

The second associated moment condition is obtained by passing expectations with respect to the ergodic measure at time t then using Tonelli's theorem to reveal a geometric progression on the right hand side.

$$\begin{aligned}(1-\alpha)^{1/(\theta-1)}\mathbb{E}\frac{u'(Y)Y}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}&\\+\alpha\beta(1-\alpha)^{1/(\theta-1)}\mathbb{E}\frac{1}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}\frac{u'(Y)Y(1+\pi)^{\theta-1}}{1-\alpha\mathbb{E}\beta(1+\pi)^{\theta-1}}&\\=\frac{\theta}{\theta-1}\mathbb{E}\frac{u'(Y)Yv'(\Delta Y/A)/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}&\end{aligned} \tag{4.7}$$

Note the right hand side is in fact the true stochastic steady state of $\aleph$ the weighted marginal cost.

For price dispersion there is just the one equation as we have assumed there are no shocks to price dispersion

$$\Delta=\mathbb{E}\Delta=\frac{1}{(1-\alpha)^{1/(\theta-1)}}\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})} \tag{4.8}$$

where we require that

$$\mathbb{E}(1+\pi)^{\theta}<1/\alpha$$

. For the standard non-stochastic trend inflation model the same condition would arise without expectation. In the overwhelmingly likely case $\theta>2$[18], the expression would be convex in inflation so that the upper bound on inflation be strictly lower

[18] corresponding to an economy in its flexible price limit would have markups of less than 100%

than in non-stochastic steady state.

$$\pi < (1/\alpha)^{1/\theta} - 1 = \bar{\pi}^{NSS}$$

Moreover, the negative covariance term would tighten the bound further. It is intuitive that incorporating volatility tightens the conditions for obtaining a stochastic equilibrium. It follows from theorem 12 that an excessively loose monetary policy where the rate of inflation is too high or too volatile could be arbitrarily costly to social welfare. Whilst, the extreme predictions are a figment of a stylized model the issue of blow up of the proposed solution is important for practical implementation. Volatility is clearly not second order in the New Keynesian world. Interpreting the interest rates associated to the fixed point of the cocycle as natural rates the Euler implies the following natural rates of interest real and nominal are

$$i(\pi, \Delta, Y) = \frac{u'(Y)\psi}{\beta \mathbb{E} u'(Y)\psi} - 1 \tag{4.9}$$

$$r(\pi, \Delta, Y) = \frac{u'(Y)\psi}{\beta(1+\pi)\mathbb{E} u'(Y)\psi} - 1 \tag{4.10}$$

These are derived from the mapping $Y \to \mathbb{E}_t Y_{t+1}$ the associated moment condition is.

$$\frac{u'(\mathbb{E}Y)\mathbb{E}\psi}{\beta \mathbb{E} u'(Y)\psi} \tag{4.11}$$

The other moment condition found by passing expectations is the following which could be estimated by a standard orthogonality condition

$$\mathbb{E}\frac{u'(Y)\psi}{1+i} = \beta \mathbb{E}\psi u'(Y) \tag{4.12}$$

where the associated value of $\int i \, \mathrm{d}\mu$ would be its stochastic steady state. The connection between the interest rate prevailing under the fixed point of the cocycle and that of its true stochastic steady state is obtained by equating the time $t+1$ terms which shows that

$$i(\pi, \Delta, Y) = \frac{u'(\mathbb{E}Y)\mathbb{E}\psi}{\mathbb{E}u'(Y)\psi/(1+i)} - 1 \tag{4.13}$$

This completes the characterization of the steady state.

4.3.1 Comparative Statics

This subsection compares the stochastic equilibrium with its non-stochastic counterpart including so called trend inflation where $\pi \neq 0$. I consider both the real members of Z, Δ and Y and other macroeconomic variables determined by Z such as wages W and labor L. It is important to remember that with these additional variables the stochastic equilibrium cannot be obtained by evaluating their value attached to the present value of the cocycle $\mathbb{E}Z$. For example the stochastic steady marginal costs $\mathbb{E}MC \neq \nu'(\Delta Y/A)/A = \nu'(\mathbb{E}\Delta\mathbb{E}Y/\mathbb{E}A)/\mathbb{E}A$ since certainty equivalence does not hold for general functions.[19] Proofs involve applications of Jensen's inequality and monotonicity conditions.

4.3.2 Non-Stochastic Steady State with Trend Inflation

Here for convenience I provide expressions for the non-stochastic steady state with trend inflation arising in the model laid out in chapter two. The non-stochastic steady state discussed there constitutes a special case, whilst any non-stochastic steady state can be viewed as the limit of its non-stochastic counterpart. Salient features for comparison with the stochastic model are drawn out. As readers may be less familiar with non-stochastic steady state with trend inflation, I construct an example with familiar functional forms that yields a closed form solution, although, this is not essential to subsequent analysis.

$$u(C) = \frac{1}{1-\sigma}C^{1-\sigma} \tag{4.14}$$

for $\sigma \neq 1$ or where $\sigma = 1$

$$u(C) = \log(C) \tag{4.15}$$

$$\nu(l) = \frac{1}{1+\eta}l^{1+\eta} \tag{4.16}$$

It is easiest to solve first for marginal costs from the price setting relation

$$MC^{NSS} = \frac{\theta-1}{\theta}\left(\frac{1-\alpha\beta(1+\pi)^{\theta}}{1-\alpha\beta(1+\pi)^{\theta-1}}\right)\left(\frac{1-\alpha}{1-\alpha(1+\pi)^{\theta-1}}\right)^{1/(\theta-1)} \tag{4.17}$$

[19] The only instances of certainty equivalence of interest here is the cocycle f itself. Prominent examples such as linear functional and higher moments of symmetric distributions do not arise or are not worthy of emphasis here.

Therefore thanks to constant returns the wage rate is

$$W^{NSS} = \frac{\theta - 1}{\theta}\Big(\frac{1 - \alpha\beta(1+\pi)^{\theta}}{1 - \alpha\beta(1+\pi)^{\theta-1}}\Big)\Big(\frac{1-\alpha}{1-\alpha(1+\pi)^{\theta-1}}\Big)A \tag{4.18}$$

Price dispersion is simply the stochastic steady state evaluated as though inflation is certain to equal its expected value

$$\Delta^{NSS} = \frac{(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}(1-\alpha(1+\pi)^{\theta})} \tag{4.19}$$

Inverting the marginal cost function and substitute in the price dispersion relationship yields

$$Y^{NSS} = \Big(\frac{\theta-1}{\theta}\Big)^{1/(\sigma+\eta)} \frac{(1-\alpha)^{(1+\eta)/(\theta-1)}(1-\alpha(1+\pi)^{\theta})^{\eta/(\sigma+\eta)}}{(1-\alpha\beta(1+\pi)^{\theta-1})^{(1+\eta\theta)/(\theta-1)}} \times \Big(\frac{1-\alpha\beta(1+\pi)^{\theta}}{1-\alpha\beta(1+\pi)^{\theta-1}}\Big)^{1/(\sigma+\eta)} A^{(1+\eta)/(\sigma+\eta)} \tag{4.20}$$

The resource constraint yields the solution for labor supply

$$l^{NSS} = \Big(\frac{\theta-1}{\theta}\Big)^{1/(\sigma+\eta)} \frac{(1-\alpha)^{(1-\sigma)/(\theta-1)(\sigma+\eta)}}{(1-\alpha(1+\pi)^{\theta})^{\sigma/(\sigma+\eta)}} \times (1-\alpha(1+\pi)^{\theta-1})^{(\sigma\theta-1)/(\theta-1)(\sigma+\eta)} \Big(\frac{1-\alpha\beta(1+\pi)^{\theta}}{1-\alpha\beta(1+\pi)^{\theta-1}}\Big)^{1/(\sigma+\eta)} A^{(1-\sigma)/(\sigma+\eta)} \tag{4.21}$$

Aggregate Profits are given by

$$\Pi^{NSS} = \Bigg(1 - \Big(\frac{\theta-1}{\theta}\Big)\Big(\frac{1-\alpha\beta(1+\pi)^{\theta}}{1-\alpha\beta(1+\pi)^{\theta-1}}\Big)\Big(\frac{1-\alpha}{1-\alpha(1+\pi)^{\theta-1}}\Big)^{1/(\theta-1)}\Bigg) Y^{NSS} \tag{4.22}$$

Note that with trend inflation not all firms will make different profits depending on when they last reset their price. Denoting the time since the last reset by the age of its price a

$$\Pi_a = \Big((1+\pi)^{-a} - MC\Big)(1+\pi)^{-a\theta}Y \tag{4.23}$$

Indeed with positive trend inflation there will be some firms making negative profits. To understand this fix $\pi > 0$ and consider the infinite limit $\lim_{a\to\infty} \Pi_a = -\infty$ and then continuity of profit with respect to a implies a cutoff such that all firms with current price spell $a \geq \bar{a}$ will make negative profits. The bracketed term is the steady state mark up. Present discounted welfare is

$$U^{NSS}=\frac{1}{1-\beta}\Bigg[\frac{1}{1-\sigma}\left(\frac{\theta-1}{\theta}\right)^{(1-\sigma)/(\sigma+\eta)}(1-\alpha)^{(1-\sigma)(1+\eta)/(\theta-1)(\sigma+\eta)}\times$$
$$\frac{(1-\alpha(1+\pi)^{\theta})^{(1-\sigma)\eta/(\sigma+\eta)}}{(1-\alpha(1+\pi)^{\theta-1})^{(1-\sigma)(\eta\theta+1)/(\theta-1)(\sigma+\eta)}}\left(\frac{1-\alpha\beta(1+\pi)^{\theta}}{1-\alpha\beta(1+\pi)^{\theta-1}}\right)^{(1-\sigma)/(\sigma+\eta)}\times$$
$$A^{(1+\eta)(1-\sigma)/(\sigma+\eta)}-\frac{1}{1+\eta}\left(\frac{\theta-1}{\theta}\right)^{(1+\eta)/(\sigma+\eta)}(1-\alpha)^{(1-\sigma)(1+\eta)/(\theta-1)(\sigma+\eta)}\times$$
$$\frac{(1-\alpha(1+\pi)^{\theta-1})^{(1+\eta)(\sigma\theta-1)/(\theta-1)(\sigma+\eta)}}{(1-\alpha(1+\pi)^{\theta-1})^{\sigma(1+\eta)/(\sigma+\eta)}}\left(\frac{1-\alpha\beta(1+\pi)^{\theta}}{1-\alpha\beta(1+\pi)^{\theta-1}}\right)^{(1+\eta)/(\sigma+\eta)}\times$$
$$A^{(1-\sigma)(1+\eta)/(\theta-1)(\sigma+\eta)}\Bigg] \quad (4.24)$$

Now I compare the stochastic steady state with its non-stochastic analog. The natural starting point is price dispersion.The following lemma will prove useful[20]

Lemma 4. *Suppose that U and V are strictly increasing functions integrable with finite mean when integrated with respect to a measure μ not concentrated on a single point and U is non-negative then $\mathbb{E}UV > \mathbb{E}U\mathbb{E}V$*

Proof.

$$\mathbb{E}UV=\int U(V-\mathbb{E}V)\mathrm{d}\mu+\mathbb{E}V\mathbb{E}U \quad (4.25)$$

By assumptions placed on the measure, the first term can be decomposed into the (non-zero) contribution of terms above and below the expected value $\mathbb{E}V$

$$\int U(V-\mathbb{E}V)\,\mathrm{d}\mu=\int_{V>\mathbb{E}V}U(V-\mathbb{E}V)\,\mathrm{d}\mu+\int_{V<\mathbb{E}V}U(V-\mathbb{E}V)\,\mathrm{d}\mu \quad (4.26)$$

(Strict) Monotonicity of U, V and the Lebesgue integral allow us to conclude that

$$\int_{V>\mathbb{E}V}U(V-\mathbb{E}V)\,\mathrm{d}\mu>U(\mathbb{E}V)\int_{V>\mathbb{E}V}(V-\mathbb{E}V)\,\mathrm{d}\mu \quad (4.27)$$

Similarly when V is below average by the same argument

$$\int_{V<\mathbb{E}V}U(\mathbb{E}V-V)\,\mathrm{d}\mu<U(\mathbb{E}V)\int_{V<\mathbb{E}V}(\mathbb{E}V-V)\,\mathrm{d}\mu$$

[20] This result is not novel to pure mathematicians it is a version of Chebyshev's algebraic inequality and is a corollary of a result on p.248 of Pecaric and Fink [1993].

now reversing yields

$$\int_{V<\mathbb{E}V} U(V-\mathbb{E}V)\,\mathrm{d}\mu > U(\mathbb{E}V)\int_{V<\mathbb{E}V}(V-\mathbb{E}V)\,\mathrm{d}\mu \tag{4.28}$$

combining the numbered equations completes the proof. □

The following theorem compares the stochastic with its non-stochastic counterpart as the inflation rate changes under the empirically highly plausible case that $\theta > 2$. A single crossing property is established via a subtle application of Jensen's inequality. I am able to prove that when inflation is positive price dispersion is increasing, convex and strictly greater than in non-stochastic steady state. Unlike, in non-stochastic steady state Δ is positively persistent and increasing at moderate rates of deflation. This point could have substantial empirical implications. Sustained periods of rapid deflation are extremely rare, whilst flat price levels with periods of shallow price decline have been common in economic history. Japan's recent economic history subscribes roughly to this narrative. Therefore my analysis calls into question the empirical significance of negative trend inflation.

Theorem 12. *There exists $\underline{\underline{\pi}} < \underline{\pi}$ such that Δ is strictly increasing (decreasing) if and only if $\pi \gtrless \underline{\pi}$ and Δ is convex (concave) above (below) if and only if $\pi \gtrless \underline{\underline{\pi}}$ and if $\theta \geq 2$ $\underline{\pi} < 0$*

Proof. Focus first on $\underline{\pi}$ we know by continuity of the first derivative that follows from smoothness that this must be a stationary point. The derivative $\mathrm{d}\Delta/\mathrm{d}\pi$ is

$$\begin{aligned}\frac{\alpha\theta}{(1-\alpha)^{1/(\theta-1)}}\Bigg[\mathbb{E}(1+\pi)^{\theta-1}\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})^2}-&\\ \frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1+\pi)^{\theta-2}}{1-\alpha\mathbb{E}(1+\pi)^{\theta}}\Bigg]&\end{aligned} \tag{4.29}$$

To prove existence of the bounds examine the limiting behavior of the derivatives around the poles. At the positive pole.

$$\bar{\pi} = \{\pi = d^{-1}(0) : \, d = 1-\alpha\mathbb{E}(1+\pi)^{\theta}\}$$

$$\lim_{\pi\to\bar{\pi}} \mathrm{d}\Delta/\mathrm{d}\pi = \infty$$

from the domination of the first term in the bracket which is $\mathbb{O}(x^{-2})$ term. Now there are more complications with the lower pole $\pi =\rightarrow -100\%$ as the limit is zero for $\theta > 2$. However, since it represents the non-negativity constraint on the price we know that this pole is shared by both the whole model and the steady state. Therefore whatever the other parameters every sequence of (stochastic equilibrium) measures must converge on a sequence of non-stochastic steady state as inflation approaches the pole. Therefore

$$\lim_{\pi\to-100\%} \mathrm{d}\Delta/\mathrm{d}\pi \to \frac{\alpha\theta}{(1-\alpha)^{1/(\theta-1)}}\frac{(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}\pi}{(1-\alpha(1+\pi)^{\theta})^2} < 0$$

in the neighborhood of the pole. The intermediate value theorem proves the existence of a stationary point. Since the derivative must move from negative to positive uniqueness will prove the existence of the boundary value $\underline{\pi}$.

Now turning to the convexity claim direct computation reveals that

$$\begin{aligned}\mathrm{d}^2\Delta/\mathrm{d}\pi^2 = {} & \frac{\alpha\theta(\theta-1)}{(1-\alpha)^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta-2}\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})^2} - \\ & \frac{\alpha^2\theta^2}{(1-\alpha)^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta-1}\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1+\pi)^{\theta-2}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})^2} + \\ & \frac{2\alpha^2\theta^2}{(1-\alpha)^{1/(\theta-1)}}(\mathbb{E}(1+\pi)^{\theta-1})^2\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})^3} - \\ & \frac{\alpha\theta(\theta-2)}{(1-\alpha)^{1/(\theta-1)}}\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1+\pi)^{\theta-3}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})} + \\ & \frac{\alpha^2\theta}{(1-\alpha)^{1/(\theta-1)}}\mathbb{E}\frac{(1+\pi)^{2(\theta-2)}}{(1-\alpha(1+\pi)^{\theta-1})^{(\theta-2)/(\theta-1)}}\frac{1}{1-\alpha\mathbb{E}(1+\pi)^{\theta}} \\ & - \frac{\alpha^2\theta^2}{(1-\alpha)^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta-1}\frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1+\pi)^{\theta-2}}{(1-\alpha\mathbb{E}(1+\pi)^{\theta})^2} \end{aligned} \tag{4.30}$$

As with the first derivative we can establish a candidate for the lower bound in this case $\underline{\underline{\pi}}$ by taking limits around the poles we can see that

$$\lim_{\pi\to\bar{\pi}} \mathrm{d}^2\Delta/\mathrm{d}\pi^2 = \infty$$

from the domination of the third term that is $\mathbb{O}(x^{-3})$. Now in the limit as we approach the lower pole, the fourth term is dominant since $\theta > 2$ so we know that

$$\lim_{\pi \to \bar{\pi}} \mathrm{d}^2\Delta/\mathrm{d}\pi^2 \to \mathrm{d}^2\Delta^{NSS}/\mathrm{d}\pi^2 < 0$$

in this case the intermediate value theorem applied to the second derivative provides a candidate stationary point which will be equal to $\underline{\underline{\pi}}$ if we can establish uniqueness. Next is the crucial step connecting convexity in stochastic equilibrium with the level of the stochastic steady state relative to its non-stochastic counterpart. By applying the Jensen-Cheyshev inequality from Proposition 5 to the stochastic steady state Δ along with iterated expectations we discover that $\Delta \gtreqless \Delta^{NSS}$ for $\mathrm{d}^2\Delta/\mathrm{d}\pi^2 \gtreqless 0$ and therefore by continuity $\Delta = \Delta^{NSS}$ if and only if $\mathrm{d}^2\Delta/\mathrm{d}\pi^2 = 0$

It is easiest to establish first the non-positivity of the bounds, starting with the first derivative. The following sequence of inequalities apply when $\theta \geq 2$.

$$\begin{aligned}
\frac{d\Delta}{d\pi} &> \frac{\alpha\theta}{1-\alpha(1+\pi)^{\theta}}\mathbb{E}\Big[(1+\pi)^{\theta-2}\Big((1+\pi)\Delta - \frac{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}}\Big)\Big] \\
&> \frac{\alpha\theta}{1-\alpha(1+\pi)^{\theta}}\mathbb{E}\Big[(1+\pi)^{\theta-2}\Big((1+\pi) - \frac{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}}\Big)\Big] \\
&\geq \frac{\alpha\theta}{1-\alpha(1+\pi)^{\theta}}\mathbb{E}(1+\pi)^{\theta-2}\Big((1+\pi) - \frac{\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}}\Big) \\
&\geq \frac{\alpha\theta}{1-\alpha(1+\pi)^{\theta}}\mathbb{E}(1+\pi)^{\theta-2}\Big((1+\pi) - \frac{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}}\Big)
\end{aligned}$$

The first, second and fourth follow from proposition 5. The first and fourth are Jensen-Chebyshev, the second uses the properties of Δ and monotonicity of the Lebesgue integral. The fourth requires $\theta \geq 2$ as the relevant second derivative changes sign (becomes convex) as it approaches the negative pole as

$$\lim_{\substack{\theta<2 \\ \pi\to-100\%}} \frac{\alpha(2-\theta)(1+\pi)^{\theta-3}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}} - \frac{\alpha^2(1+\pi)^{2(\theta-2)}}{(1-\alpha(1+\pi)^{\theta-1})^{(\theta-2)/(\theta-1)}} = \infty$$

The third inequality follows from lemma 5 since with $\theta > 2$ both $(1+\pi)^{\theta-2}$ and the bracketed term inside the expectation are strictly increasing as the first derivative of the brackets is

$$1 + \alpha(1+\pi)^{\theta-2}/(1-\alpha(1+\pi)^{\theta-1})^{(\theta-2)/(\theta-1)}(1-\alpha)^{1/(\theta-1)} > 0$$

the weak inequality incorporate the case $\theta = 2$ where $(1+\pi)^{\theta-2}$ is a constant rather than strictly increasing. Therefore the root $\pi = 0$ is unique and the strict inequality implies that any $\underline{\pi} < 0$.

Turning to the convexity bound in search of a contradiction. It is necessary to work first with the non-stochastic steady state Δ^{NSS} its second derivative is simply (5.32) with the expectations removed. It can be written as follows

$$\begin{aligned}(1+\pi)^{-1}\left(\frac{(\theta-1)+\alpha(\theta+1)(1+\pi)^{\theta}}{1-\alpha(1+\pi)^{\theta}}\right)\frac{\mathrm{d}\Delta}{\mathrm{d}\pi}+ \\ \frac{\alpha\theta}{(1-\alpha)^{1/(\theta-1)}}\frac{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{1-\alpha(1+\pi)^{\theta}}(1+\pi)^{\theta-3}+ \\ \frac{\alpha^2\theta}{(1-\alpha)^{1/(\theta-1)}}\frac{(1+\pi)^{2(\theta-2)}}{(1-\alpha(1+\pi)^{\theta-1})^{(\theta-2)/(\theta-1)}}\frac{1}{1-\alpha(1+\pi)^{\theta}}\end{aligned} \tag{4.31}$$

where the first term comes from the first and fourth terms of the previous expression, the second comes from combining terms two, three and six whilst the final term is term five. Since we have already established that the first derivative is (strictly) positive for (strictly) positive inflation (in Chapter 2) then we know that is strictly convex for non-negative inflation.

Furthermore, we know that Δ must be convex at the origin since it always lie strictly above $\Delta^{NSS} = 1$. This rules out $\underline{\underline{\pi}} > 0$ since by the intermediate value theorem there must be some π in the neighborhood of the origin where Δ is still convex. However to prove we can still have $\underline{\underline{\pi}} < 0$ it is necessary to rule out the possibility that Δ is convex at any positive π. Suppose this were the case there would have to be $0 < \tilde{\pi} < x$ such that Δ is strictly convex for $0 < \pi < \tilde{\pi}$. Therefore Δ^{NSS} must cut Δ from below hence at $\tilde{\pi}$ the non-stochastic steady state profile must be steeper

$$\frac{\mathrm{d}\Delta^{NSS}}{\mathrm{d}\pi} > \frac{d\Delta}{d\pi}$$

since this inequality is strict continuity ensures there is some region to the right of $\tilde{\pi}$ such that $\Delta^{NSS} > \Delta$. Since we have established Δ must become convex eventually there must be some point $\tilde{\tilde{\pi}}$ where the curve again becomes convex such that Δ is convex on $(\tilde{\pi}, \tilde{\tilde{\pi}})$. However since Δ^{NSS} is still convex while Δ is concave and both cases are strict near $\tilde{\pi}$ Gronwall's differential inequality see Pachpatte [1997] tells us that

$$\Delta^{NSS}(\tilde{\tilde{\pi}}) > \Delta(\tilde{\pi}) + (\tilde{\tilde{\pi}} - \tilde{\pi})\frac{\mathrm{d}\Delta^{NSS}}{\mathrm{d}\pi}|_{\pi=\tilde{\pi}}$$

$$\Delta(\tilde{\tilde{\pi}}) < \Delta(\tilde{\pi}) + (\tilde{\tilde{\pi}} - \tilde{\pi})\frac{\mathrm{d}\Delta}{\mathrm{d}\pi}|_{\pi=\tilde{\pi}}$$

Hence

$$\Delta^{NSS}(\tilde{\tilde{\pi}}) > \Delta(\tilde{\tilde{\pi}})$$

However by Jensen's inequality

$$\Delta^{NSS}(\tilde{\tilde{\pi}}) \leq \Delta(\tilde{\tilde{\pi}})$$

and we have reached a contradiction. Therefore Δ must be strictly convex in π when inflation is non-negative so any $\underline{\underline{\pi}} < 0$. The bulk of the work for the case of negative inflation is achieved by the following powerful single crossing result. Suppose towards a contradiction that there is we know that there is an intersection $\Delta(\tilde{\pi}) = \Delta^{NSS}(\tilde{\pi})$ where Δ is increasing but Δ^{NSS} is strictly decreasing. In search of a contradiction consider boundary of the concave region containing $\tilde{\pi}$ which can be denoted by $\tilde{\tilde{\pi}}$. This is either the negative pole or another solution to $\mathrm{d}^2\Delta/\mathrm{d}pi^2 = 0$ an application of Gronwall's inequality reveals that

$$\lim_{\pi \to \tilde{\tilde{\pi}}} \Delta^{NSS} > \Delta(\tilde{\pi})$$

whilst

$$\lim_{\pi \to \tilde{\tilde{\pi}}} \Delta < \Delta(\tilde{\pi}) - (\tilde{\pi} - \tilde{\tilde{\pi}})\frac{\mathrm{d}\Delta}{\mathrm{d}\pi}|\pi = \tilde{\pi}$$

Hence

$$\lim_{\pi \to \tilde{\tilde{\pi}}} \Delta^{NSS} > \lim_{\pi \to \tilde{\tilde{\pi}}} \Delta$$

which contradicts either Jensen's equality under continuity assumptions or convergence in distribution established at the negative pole. This establishes that any $\underline{\underline{\pi}} < \underline{\pi}$, it proves that there is a rate of inflation below which Δ is weakly concave. Therefore using the fact that inflation is strictly decreasing in the limit approaching the negative pole completes the proof of the existence of $\underline{\pi}$.

□

Proposition 9. *Provided that $\theta \geq 2$, $\pi > \underline{\underline{\pi}}$ and $\sigma \geq 1$ the stochastic steady state output $Y < Y^{NSS}$*

Proof. Take the optimal reset price condition with the object of interest the right hand side $\aleph(Z)/\beth(Z)$. The first step is to show this is a convex function of π. The easiest way to do so is to differentiate the left hand side which is a function only of π. The task is accomplished by computing the second derivative and using the parameter restriction for θ

$$\alpha(1-\alpha)^{1/(\theta-1)}(1+\pi)^{\theta-3}\left[\frac{(\theta-1)(\theta-2)(1-\alpha(1+\pi)^{\theta-1})+\alpha\theta(1+\pi)^{\theta-1}}{(1-\alpha(1+\pi)^{\theta-1})^{(2\theta-1)/(\theta-1)}}\right] > 0$$

Note that the condition is sharp in θ by considering the limits respectively as $\alpha \to 0$ and $\pi \to -100\%$. Now Jensen-Chebyshev inequality implies that left hand side strictly exceeds the right hand side evaluated at the non-stochastic steady state. This implies that Δ or Y must differ from their non-stochastic steady state values to restore equilibrium. The right hand side is clearly increasing in Δ. Therefore to correct the imbalance Δ would have to fall. However, this induces a contradiction since Δ is convex in π for the relevant parametization. Hence Y must change to equilibriate the system. To see that it must fall take the derivative of the right hand side

$$\begin{aligned}\Bigg(\frac{\nu'(\Delta Y/A)}{A}+\frac{Y\nu''(\Delta Y/A)}{A}+\alpha\beta\frac{\mathbb{E}\{\nu'(\Delta Y/A)+Y\nu''(\Delta Y/A)\}(1+\pi)^{\theta}/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Bigg)/ \\ \left(u'(Y)Y+\alpha\beta\frac{\mathbb{E}u'(Y)Y(1+\pi)^{\theta-1}}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\right)+ \\ (\sigma-1)\left(u'(Y)+\alpha\beta\frac{\mathbb{E}u'(Y)(1+\pi)^{\theta-1}}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\right)\times \\ \left(\frac{Y\nu'(\Delta Y/A)}{A}+\alpha\beta\frac{\mathbb{E}Y\nu'(\Delta Y/A)(1+\pi)^{\theta}/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right)/ \\ \left(u'(Y)Y+\alpha\beta\frac{\mathbb{E}u'(Y)Y(1+\pi)^{\theta-1}}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\right)^2\end{aligned} \tag{4.32}$$

since when $\sigma \geq 1$ the second term is non-negative and therefore unable to cancel the strictly positive first term. Sharpness depends on parametric assumptions. Suppose $0 < \sigma < 1$ and consider the joint limit where the economy approaches any non-stochastic steady state and $A \to 0$. There are two competing terms the first represented by the top two lines above is always positive and the second is negative by assumption. The idea is to show that in this limiting case the second term will dominate sufficiently close to these limiting cases. To achieve this consider the ratio of the negative term to the positive term. The proof will be complete if its limit is greater than unity. The expression is the following product

$$\lim_{\substack{A\to 0\\ \pi\to\bar{\pi}}} \left(u'(Y) + \alpha\beta\frac{\mathbb{E}u'(Y)(1+\pi)^{\theta-1}}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\right) \Big/ \left(u'(Y)Y + \alpha\beta\frac{\mathbb{E}u'(Y)Y(1+\pi)^{\theta-1}}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\right)$$
$$\times\left(\frac{Y\nu'(\Delta Y/A)}{A} + \alpha\beta\frac{\mathbb{E}Y\nu'(\Delta Y/A)/A(1+\pi)^{\theta}}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right)\Big/$$
$$\left(\frac{\nu'(\Delta Y/A)}{A} + \frac{Y\nu''(\Delta Y/A)}{A} + \alpha\beta\frac{\mathbb{E}\{\nu'(\Delta Y/A) + Y\nu''(\Delta Y/A)\}(1+\pi)^{\theta}/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right) \tag{4.33}$$

With functional forms from the previous subsection in place this can be computed with big $\mathcal{O}$ notation. The first ratio (top line) is
$\mathcal{O}(Y^{-1}) = \mathcal{O}(A^{-(1+\gamma)/(\sigma+\gamma)})$ The second term is more complicated the numerator the second line is $\mathcal{O}(\Delta Y/A)/A = \mathcal{O}(A^{-\sigma(1+\gamma)/(\sigma+\gamma)})$. Now turning to the denominator line three there are terms in this order and others in $\mathcal{O}(Y\nu''(\Delta Y/A)/A) = \mathcal{O}(A^{-(1+\sigma\gamma)/(\sigma+\gamma)})$ since $\sigma < 1$ this term will dominate and the total order of the second term will be $\mathcal{O}(A^{(1-\sigma)/(\sigma+\gamma)})$. Hence the order of the whole limit will be $\mathcal{O}(A^{-(1+\sigma\gamma)/(\sigma+\gamma)} + (1-\sigma)/(\sigma+\gamma)) = \mathcal{O}(A^{-1})$ hence the limit will be infinite. This ensures that for productivity sufficiently low and inflation sufficiently stable the derivative will be negative and the result overturned for any $0 < \sigma < 1$. This proves that $\sigma \geq 1$ is a sharp restriction. □

The empirical restrictions are not overly burdensome. To my knowledge the lowest estimate for θ is 2.5 (associated with an average mark up of 67%.) This comes from De Loecker and Eeckhout [2017] and pertains only to recent US experience. For the early part of the sample (pre 1980) however, markups averaged only around 20% corresponding to a much higher $\theta = 6$. Moreover. this appears to be a US specific trend Barkai [2017], Gutierrez [2017], driven entirely by rising dispersion across firms that is not a focus here- see Autor et al. [2017] for further discussion. All estimates in Christiano et al. [2005] are above 4.5 (see equations 3 and 4 and table 2) and equation 1 and are significantly different from 2.5 at standard levels. This is typical for the DSGE literature.

For σ microeconomic studies reveal an average estimate close to 2- Havránek et al. [2015]. Higher values make it easier to hit moments of financial data Braun and Nakajima [2011], improving our ability to explain results such as the bond premium and forward guidance puzzles. A literature has attempted to justify this with various departures from expected utility maximization and full information abstracted from here. Even if these arguments were rejected. $\sigma = 1$ at the end of our range would appear focal as the sole value that supports a balance growth path- consistent with the near ubiquitous empirical strategy of filtering away the effects of long-run growth. This comparison can be extended to other variables of distributional significance for firms and workers. The main work is done by the following application of Jensen's inequality to marginal costs.

Proposition 10. *$MC > MC^{NSS}$ if $\sigma \geq 1$ and $\pi \geq \underline{\pi}$.*

Proof. From the labor supply condition $MC = \mathbb{E}MC = \mathbb{E}(\nu'(\Delta Y/A)/A) \times (1/\phi u'(Y))$. The left hand function in the product is strictly increasing and strictly convex in (Δ, Y, $1/A$). Likewise the second function is also strictly increasing and convex in Y since $(1/\phi u'(Y))' = \sigma Y/u'(Y)$ and $(1/u'(Y))'' = \sigma(1+\sigma Y)/\phi u'(Y)$ and weakly convex in $1/\phi$. It is also strictly increasing and weakly convex in $1/\phi$. Now we know that the product of two function that are (strictly) convex and (strictly) increasing is also (strictly) increasing and (strictly) convex, hence we can apply proposition 5 with respect to (Δ, y, $1/A$). To complete the proof note that we established earlier in this chapter that for $\pi \geq \underline{\pi}$ Δ is strictly convex and increasing. Therefore MC is convex in the full set of stochastic process of the model $(\pi, Y.1/\phi, 1/A)$ and since it is strictly convex in (π, Y, $1/A$) we can apply Proposition 5 to reveal that

$$MC > \nu'(\Delta Y \mathbb{E}(1/A))\mathbb{E}(1/A)\mathbb{E}(1/\phi)/u'(Y)$$

Now note that $1/x$ is strictly convex for the case of interest $x > 0$ apply this result to the shock terms yields

$$MC > \nu'(\Delta Y/A)/A\phi u'(Y)$$

□

This result naturally extends to wages, labor supply, and profits

Proposition 11. *If $\sigma \geq 1$, $\theta > 2$ and $\pi \geq \underline{\pi}$ then $l > l^{NSS}$, $W > W^{NSS}$ and $\Pi < \Pi^{NSS}$.*

Proof. To establish the first result all we need is to show that there is a positive relationship between wages and marginal costs in stochastic steady state this follows because $\frac{\partial W}{\partial MC} = 1/\mathbb{E}(1/A) > 0$ likewise for marginal cost $\frac{\partial l}{\partial MC} = 1/\mathbb{E}(\nu_l(l)/A) > 0$ so higher marginal cost implies higher labor supply. Finally, turning to profits from the restriction on θ we know from previous results that $Y < Y^{NSS}$ then $\frac{\partial Wl}{\partial l} = \mathbb{E}W > 0$ and $\frac{\partial Wl}{\partial W} = \mathbb{E}l > 0$ imply that profits must fall to balance out the stochastic steady state budget constraint $Y = \Pi + \mathbb{E}WL$ □

With empirical eyes it is counter-intuitive that the economy becomes less efficient and the representative household worse-off because volatility causes there to be too much work and too little profit. It would seem particularly suprising that inefficiency linked to sticky prices and volatility increases both wages and hours. However, this is a figment of several stylized features of the model. The first is an efficient labor market with interior labor supply. If we introduced a frictional labor market with unemployment this result may well be reversed. Introducing additional factors of production could have similar effects if stochasticity caused the capital stock to fall-ceterus parebus the average marginal product of labor should decline also implying a reduction in labor demand. Finally, the centrality of falling profits to worker welfare could be removed by replacing the representative household with a heterogeneous agent framework as has been analyzed by Broer et al. [2016], Auclert et al. [2017], Kaplan et al. [2018] and Auclert and Rognlie [2018] among others. Moreover, I feel the comparative statics and linearization techniques presented here may benefit the heterogeneous agent research agendas predominant across much of macroeconomics, although I leave this for future researchers.

This section has compared the stochastic steady state with its non-stochastic counterpart and shown that under plausible parameter restrictions output is lower whilst distortions are higher. While there are differences in behavior around zero inflation-where the stochastic steady state maintains properties of the non-stochastic system with positive trend inflation; I believe these results could easily be extended by undertaking simulations or making additional restrictions on the tails of distributions.

4.4 Theoretical Significance and Extensions

It is worth dwelling on the theoretical significance of the equilibrium concept I postulate here, along with possible extensions of this research agenda beyond the present boundary of New Keynesian DSGE models. Equilibrium is an almost essential organizing concept in economics. Unlike its not non-stochastic counterpart stochastic equilibrium specifies all probabilistic fluctuations. It is a complete description of the business cycle in the way that a non-stochastic model approximated in deviation for and augmented with ad hoc shocks is not.

Furthermore, stochastic equilibrium encompasses the conception of non-stochastic steady state through point measures. Indeed, it can encompass many dynamic models that are not thought of as having steady states, since in a stochastic environment even plausible chaotic models have ergodic measures as shown in Lasota and Mackey [1998] and Arnold [2013]. In fact, the formulation of stochastic equilibrium here is the broadest possible equilibrium concept for which we can apply the scientific method with confidence. This is because it is defined in terms of the *ergodic distribution* the base of the ergodic hierarchy explained in Frigg et al. [2016]. It is the weakest restriction we can place on the data and still be sure of learning its statistical properties with a sufficiently large sample. Without ergodicity we either could not construct statistics to test particular hypotheses or those tests would be inconsistent.

Indeed, the definitions presented here clarify what an equilibrium is and is not. Since Walras [1874] and Marshall [1890] economists have associated an equilibrium as a point of equality between supply and demand. Walras' law explained how this followed naturally from rational decision-making by non-satiated consumers. This was continued through the aggregate demand- aggregate supply paradigm that formed the backbone of the initial synthesis between New Classical and Keynesian economics see Hicks [1937] and Samuelson [1951]. It is maintained here because firms meet all demand at the posted price. Hence, the equilibrium of this and every other variant of the New Keynesian model is *Walrasian.*

However, a problem arises when trying to incorporate nominal rigidity. A natural synthesis is to introduce nominal rigidity through the presence of fixed wages and / or prices whilst maintaining the equilibrium concept closest to classical economics; the non-stochastic steady state when there is zero inflation the simplest infinite horizon extension of Walrasian equilibrium. The problem is that in a non-stochastic price equilibrium price rigidity will never bind- firms will never want to change their price.

In stochastic equilibrium, by contrast price rigidity always bites there is a distribution of prices and there will always be some firms with sub-optimal prices. Adjustment is part of the equilibrium with idiosyncratic shocks even if the economy is in equilibrium we can expect and in fact be (almost) sure that some firms prices will move further from optimal levels others will move closer. However, we can be (almost) sure that the two will cancel out.

Stochastic equilibrium itself is not completely novel to economists as discussed earlier it has been explored informally in financial economics, where there is portfolio allocation under uncertainty. What is surely novel however is the divorce between the conception of equilibrium and optimization. Walrasian equilibrium is a consequence of static optimization, whilst most modern economics macro and micro features the study of equilibrium associated with agents solving dynamic infinite horizon optimization problems. This has caused confusion amongst macroeconomists it is not uncommon to read statements like

> "An equilibrium can now be defined as a collection of stochastic processes ... given a path for the path for the exogenous variables ... and the initial conditions ... that satisfy equations ... "

This is in fact a *trajectory.* [21]

Indeed, optimization alone cannot characterize the equilibrium or other behavior of

[21] The quote comes from p5-6 of the seminal paper Eggertsson [2010]. I could have picked many similar examples.

a benchmark New Keynesian model because the model's central feature is that some firms are not displaying optimal behavior at any given point in time. This is a radical methodological departure from previous ideas of economic equilibrium. It answers to the most serious and widespread criticism leveled at the economics profession that we rely excessively on optimization. Behavioral economics pursues a complementary approach in loosening the straitjacket of rational maximizing behavior. However, to my knowledge it always does so in a framework featuring optimization in all periods.[22] Complementarity between the two approaches ought to be explored.

By contrast, the heterodox approach where formal methods have been used has tended to avoid explicit stochastic formulations reflecting perhaps perceived associations with rational expectations. For example Homburg [2017] exposits a DGE (Dynamic General Equilibrium) framework based on the temporary equilibrium approach where a shock generates a sequence of moving equilibrium, see also Grandmont [2006]. The mathematical justification for this result is empirically problematic. Schlicht [1985] and Schlicht [1997] provide details. It requires a clear distinction between fast and slow moving variables. Economics cannot provide this. Stickiness and delayed adjustment is ubiquitous in macroeconomic datasets.[23] DGE has the unfortunate distinction of missing both strands of Keynesian analysis uncertainty and price rigidity. Temporary equilibrium is actually a classical idea first explored by Marshall [1890]. He was working in a partial equilibrium context where he assumed price was the fast moving variable whilst capacity was slow moving. Hahn et al. [1961] extended his analysis into general equilibrium. He was fully aware of the disjuncture between espousing Keynesian macroeconomics based around nominal rigidity and Walrasian microeconomics based on extreme flexibility a point acknowledged in Hahn [1977].

Stochastic equilibrium addresses this problem whilst incorporating heterodox themes. It offers a powerful opportunity to bring heterodox economists into the mainstream economics fold. Up to now economics has had too narrow a conception of equilibrium. It has hindered our understanding of macroeconomics. A broad scientific consensus

[22] Typically in behavioral economics a benchmark rational optimizing model is modified by either changing the information set of the agents so that they have "wrong beliefs" or changing the objective function they maximize so that they have "different objectives" from the consequentialist utilitarian benchmark. All the formal models in Wilkinson and Klaes [2012] can be viewed through this lens.

[23] This approach is called scale separation in physics. Applications have been more successful since distinctions between fast and slow moving variables are clearer Kokotovic et al. [1980].

can be reached.

Moreover, stochastic equilibrium is a technical innovation with powerful implications for econometrics and microeconomics. Standard estimators are not applicable to the stochastic equilibrium. The dependence of the support of inflation on the parameters, crucial to the proof of Theorem 12, makes the maximum likelihood estimator inconsistent in general. For exposition consult Greene [2003] and Hirano and Porter [2003]. [24] The blow up of the model at the boundary means asymptotic normality is lost for a general class of extremum estimators including Bayesian and Generalized Method of Moment estimators see Newey and McFadden [1994]. It also means that the Cramer-Rao and alternative lower bounds on other unbiased estimators also become unachievable Yaakov et al. [2014] and Lu et al. [2017]. The novel challenge faced by all potential estimators is that the space of admissible parameters is itself unknown, whereas usually there is a given parameter set from which the maximum is chosen. In practical applications one must guard against an estimator yielding a non-admissible parameter configuration- which could occur if the parameter set were not estimated sufficiently accurately.

Since stochastic equilibrium is defined by moment conditions (six in this section- two for each of the Phillips, price dispersion and Euler equations) the natural estimator would be GMM. Indeed, in such a non-linear model where higher moments can effect the existence of equilibrium there would be great concern about the imposition of particular distributions on error with largely ad hoc justification. However, only two of the six conditions (the basic Euler and price dispersion evolution) yield a sample moment condition $\sum_T g(Y_T, \theta|X_T) = 0$ This is because the other four conditions feature non-linear relationships among expectations of functions. The solution to this problem lies beyond the scope of the thesis.

Any solution is likely to involve application of techniques from optimal transport. Optimal transportation is a branch of mathematics that studies how to transform one probability distribution into another whilst minimizing a cost function that depends on the probability mass moved. Its salient feature is the optimization of a joint

[24] This result is well understood in the econometrics. There is an econometric literature investigating the non-standard limiting distribution commonly associated with models such as search and auctions where maximizing behavior connects the support and parameters see for example Flinn et al. [1982], Smith [1985], Christensen and Kiefer [1991], Donald and Paarsch [2002] and Chernozhukov et al. [2004].

distribution subject to restrictions on the marginals. It is an active area of research across mathematics physics and economics. It has yielded one Fields medal (Cedric Villani- 2010) and a joint Nobel prize (Leonid Kantorovich and Tiajing Koopmans- 1975). Optimal transport is not novel to the econometric toolkit. It has proven useful in the discrete choice framework where it is used to map between market shares and unobserved consumer preferences Galichon and Salanie [2015] and Chiong et al. [2016]. The crucial technical innovation has been multivariate extension of quantile regression- Koenker and Bassett Jr [1978], by Carlier et al. [2016] and Chernozhukov et al. [2017] which has been used to identify multivariate risk in Ekeland et al. [2012] and generalize partial identification approaches from Matzkin et al. [2003] and Heckman et al. [2010] to the multidimensional case.[25] Galichon [2017] provides a concise review of this new econometric literature. Galichon [2016] is a monograph that explains current economic applications with a strong emphasis on computational implementation. Villani [2008] is the seminal mathematical reference.

The most direct crossover with microeconomics is equilibrium refinement. Nash equilibrium often gives rise to multiple solutions. It is often a case that one or more seems more intuitive than others. The idea is to use a stronger definition of Nash equilibrium to rule out certain equilbria. The most famous sub-game perfection rules out time inconsistent threats. A popular class of refinements involve perturbations to the players strategies in an attempt to remove equilibria that are not robust to small noise or deviations from rationality for example the following equilibrium conceptsSelten [1975] (trembling hand), Myerson [1978] (proper),Kreps et al. [1982] (sequential) and van Damme [1987] (approachable). Recent extensions such as Fudenberg and He [2017b], Fudenberg and He [2017a] and Milgrom and Mollner [2018] have developed more powerful refinements based around the limits of rational learning and costly review.

[25] Connections with microeconomic theory have also been uncovered often in areas adjoining the econometric applications. For example in matching, pricing and bargaining problems following the seminal contribution of Chiappori et al. [2010].The paper proved that classical allocation problems such as those studied by Shapley and Shubik [1971] andBecker et al. [1973] are equivalent to a classical optimal transport problem. Optimal transport has allowed existence and uniqueness results in multidimensional matching problems and allowed the removal of restrictions such as quasi-linearity Figalli et al. [2011], Decker et al. [2013], Dupuy and Galichon [2014], Chiappori et al. [2016], Noldeke et al. [2017] and Greinecker and Kah [2018]. Likewise Pass [2012], McCann and Zhang [2017], Zhang [2017] and Warren [2017] have used optimal transport to attack problems of multi-good monopoly. Blanchet and Carlier [2014] and Blanchet and Carlier [2016] combine mean field games and optimal transport in an extension of Cournet competition.

Rigidity in the sense of players moving at random frequencies as in Calvo would be a natural basis to build a refinement applicable to infinite horizon games. It could allow game theorists to test the significance of order of actions and in particular the common assumption of simultaneous play. The nearest links in the recent literature are with the prominent infinite horizon refinement paper Simon and Stinchcombe [1995] and the rational inattention literature, such as Ravid [2017]. Stochastic equilibrium here may provide technical spillovers for example to the stochastic games literature as well as a direct route to comparative statics without the need to adopt continuous time scale, Sannikov [2012] and Simon [2016] provide background surveys.

Stochastic equilibrium promises to be a unifying theme across economics, connecting more tightly than ever before theory and empirics and the three traditional subdisciplines microeconomics, macroeconomics and econometrics. There is likely to be increased demand for mathematical as well as computational skills. Growing technical requirements may precipitate greater specialization in training and research within the profession. The impact of stochastic equilibrium research promises to be both multifaceted and profound.

Chapter 5

Valid and Invalid Approximation

This chapter provides concrete explanation about what is wrong from a dynamical systems perspective with the current practice of approximating the New Keynesian Phillips curve about its zero inflation non-stochastic steady state, in particular by showing which properties prevent one applying various versions of the Grobman-Hartman theorem used to justify linear approximation. It finishes by showing that linear approximation carried out local to the stochastic steady state constitutes a valid description of local dynamics.

5.1 Bifurcation, Valid Approximation and Non-negativity

A bifurcation occurs where small changes in the parameter values give rise to abrupt changes in the "qualitative" or topological behavior of the system. For a suitably smooth dynamical system such as ours this can only occur at a fixed point and is associated with changes in the stability, existence or uniqueness of equilibrium. Consider a sequence of stochastic equilibrium that passes the non-stochastic zero inflation steady state as Δ decreases while $\Delta > 1$ theorem 11 assures a stochastic equilibrium exists, when $\Delta = 1$ we have the familiar zero inflation non-stochastic steady state. However as soon as $\Delta < 1$ the equilibrium disappears (proposition 5). A change in the number of equilibria is called a steady state bifurcation, as we are only changing one parameter Δ the average value of price dispersion it is called a codimension one bifurcation. Therefore at the zero inflation non-stochastic steady state there is a **Local Codimension One Steady State Bifurcation**.

The formal arguments run as follows

Definition 20. A dynamical system $\mathbb{E}Z_{t+1} = f(Z_t, \gamma, \mathcal{U}_t^S)$ has a bifurcation at $\gamma^0 \in \Gamma$ if for every neighborhood of γ^0 such that $f(Z^*(\gamma^0), \gamma, \mathcal{U}_t^S)$ are not topologically equivalent.

Proposition 12. *The Calvo model laid out in Chapter 2 has a bifurcation at its zero inflation non-stochastic steady state*

Proof. From (2.23) the marginal cost relationship and (2.42) the marginal cost Phillips curve we know that the dynamics are independent of lagged inflation at the zero inflation non-stochastic steady state. However Proposition 1 tells us that in general (almost everywhere by Sard's theorem) the dynamics depend on π_{t-1}. Therefore the non-stochastic steady state is not topologically conjugate to any of its neighborhoods. □

Note that the bifurcation has to do with past inflation not price dispersion. The correct approximation at the zero inflation non-stochastic steady state depends on deviations in (y_t, π_t, Δ_t) even at this zero inflation non-stochastic steady state setting $\hat{\Delta}_t = \hat{\pi}_t$ the correct relationship is $\mathbb{E}_t\hat{\Delta}_t = \alpha\Delta_t$ which I established would be non-zero for $\pi_t \neq 0$. If we had sufficient data the correct model of the cocycle to estimate would need to feature (y_t, π_t, Δ_t).

Definition 21. A steady state bifurcation occurs where the local topological equivalence between the steady state and the parameter breaks down or the stability of some equilibria changes.

This definition is an extension of previous results designed to incorporate the case here where eigenvalues are not used to define the

Proposition 13. *There is a steady state bifurcation in π parameter at ZINSS.*

Proof. Focus on the behavior of price dispersion in response to perturbations in $\pi = 0$ and $\Delta = 1$

$$\Delta(1, \pi) = (1 - \alpha(1 + \pi)^{\theta-1})^{1/(\theta-1)} + \alpha(1 + \pi)^{\theta-1}$$

This is a smooth function of π and differentiation reveals that

$$\frac{\mathrm{d}\Delta(1, \pi)}{\mathrm{d}\pi} = \alpha(\theta - 1)(1 + \pi)^{\theta-2}[(1 + \pi) - \frac{(1 - \alpha(1 + \pi)^{\theta-1})^{1/(\theta-1)}}{(1 - \alpha)^{1/(\theta-1)}}]$$

which we established before is zero if and only if $\pi = 0$ and the inverse function theorem completes the proof. □

The final result verifies that the codimension is unity and that we do indeed have a steady state bifurcation. The economic intuition behind this mathematical argument is a policy experiment where a so called MIT shock is taking place at the ZINSS. A bifurcation is occurring because this is the place where the the impulse response to price dispersion changes sign. This reflects the fact that there is a strict minimum value to Δ covered in proposition 5.

An instructive alternative is to work with the stochastic steady state and the dispersion parameters of its errors. Proposition 5 and the assumptions made on the errors in chapter two that ensure that shocks canceling out is improbable allows us to conclude $\Delta^{SS} = 1$ if and only if all shocks have zero dispersion. Therefore the zero inflation non-stochastic steady state could be viewed as a bifurcation of the stochastic model in the dispersion parameters of the errors[1]. For the model to be estimable this codimension would have to be greater than unity and could be arbitrarily large depending on the detail with which we wanted to describe the error process. This time the non-negativity constraint is clearer as it pertains to the parameter itself.

This perspective is insightful. with non-negativity in place we only need consider perturbations where $\gamma > 0$ where we know $\Delta > 1$. This implies the fixed point is repulsive and the dynamics around ZINSS are therefore equivalent to a **supercritical pitchfork bifurcation** where a structurally stable steady state becomes unstable- see Crawford [1991] p 1002 & 1005 or for more details of the stochastic case consult Arnold [2013] chapter 9. Refinements through equilibrium stability have been popular such as the E-stability method that uses expectation shocks to determine whether models are learnable and therefore robust to plausible uncertainty. By inclusion of expectation shocks into γ we can see that in fact the ZINSS is not in fact expectation stable. This overturns results that suggest that close to the ZINSS the New Keynesian model is free from these problems- see Evans and Honkapohja [2001].

The properties of the bifurcation at ZINSS contrast with those widely studied in physics or other disciplines. Central to the disjuncture is differing ambient space.

[1] Consult Arnold [2013] for a definition more general than required here and discussion of stochastic bifurcation.

With models of the physical world it is natural to work with $\mathbb{R}^n$ or more generally suitable smooth manifolds- see for example Crawford [1991] or Kuznetsov [2013]. Negative axes can represent differing directions or charges, They are integral to studying symmetry properties of equilibria and describing conservation laws. In economics however the emphasis is on optimization problems where non-negativity constraints are important. In the case of Δ this forces us to work with the half-open interval $[1.\infty)$ which is not a manifold it is not homeomorphic to $\mathbb{R}$. The ZINSS is on this boundary where local dynamics are non-Euclidean.

Its salient features are a strictly convex relationship between a state variable Z_i and a perturbation parameter γ local to a unique steady state $Z^*(\gamma^0)$. This ensures that $Z \geq Z_i^*$- for proof suppose the converse the intermediate value theorem to contradict uniqueness. This is not possible if the state space is a manifold we would never observe $Z < Z_i^*$. This creates novel stability properties. On $\mathbb{R}^n$ bifurcations occur where structural stability properties change. This is not the case here. The steady state is structurally stable on both sides of the bifurcation, its local stability properties are unaltered. Therefore eigenvalues ofvalid linear approximation will not lie on the unit circle. To further appreciate the source of the bifurcation it is instructive to compare the model with a common manifold bifurcation called a saddle node.

Bifurcations are often associated with changes in the number of equilibria for example the famous saddle node where two fixed points coalesce. This interpretation can be restored if we think of deviations in Δ the steady state rather than the parameter π. If we plot $\Delta - 1$ on π and perturb the curve vertically downwards we can see that for a positive perturbation equivalent to relaxing the lower bound on price dispersion to $\Delta^* < 1$ the two "equilibria" emerge associated with the two branches of the quadratic. Whilst if the upper bound is raised to $\Delta^* > 1$ then there is no "equilibrium" anymore. Therefore the dynamics are mimicking a **saddle-node bifurcation** in π with parameter $\Delta^* = 1$ when we consider negative perturbations about the lower bound on price dispersion,whereas the disappearance of the equilibrium when we raise the lower bound cannot be observed for recurrent self-map on a manifold where the intermediate value theorem guarantees a fixed point. This subject is ripe for further mathematical investigation.

Note that this bifurcation does not arise in the Rotemberg model where price dispersion is absent. This has significant empirical content. It means that results obtained from studies using the 'incorrect' ZINSS approximation to the Calvo New Keynesian model can be reinterpreted from the same model with Rotemberg pricing thanks to Proposition 4. This bifurcation is a property of the price aggregator and models with non-optimal price setting- therefore it extends to Generalizations of the Calvo model such as those with heterogeneity in repricing frequency and to models with Taylor contracts- where there is no equivalence with Rotemberg to fall back on. I feel my brief study here improves upon previously literature Benhabib et al. [2002], Barnett et al. [2002], Barnett and Duzhak [2010], Stachurski et al. [2012], La'O [2013], Beaudry et al. [2015] and Brito et al. [2017] are part of a literature that applies similar techniques to physics often relying on specific timing conditions, restrictions on parameters or admissible uncertainty. Many applied economists seem to have been aware of bifurcation associated with non-negativity constraints in finance, monetary policy and investment allocation. Although, to my knowledge the only paper to actively mention bifurcation in the context of ZINSS is Kim et al. [2011] where it is used to motivate high order approximation[2] I feel this section puts economics on a firmer footing in our understanding of where to take first order approximations and the potential pitfalls involved.

5.2 Hyperbolicity

The following section bridges this chapter. It links bifurcation to the notion of hyperbolicity that will be used later to derive a notion of valid approximation. A system displays *hyperbolic dynamics* if there are expanding and contracting directions for the derivative under the defining operator T. Its significance is that fixed points of hyperbolic systems possess powerful stability and representation properties, that are desirable for economic analysis, which I discuss later in this Chapter. Thus far it has been assumed that the basic New Keynesian model displays these properties local to its zero inflation non-stochastic steady state. Here I show it does not.

First some preliminaries, the set of all tangents $\mathbf{T_X}$ to T is called its *tangent bundle*.

[2] Judd [1998] also discusses such techniques in his famous numerical analysis textbook aimed at economists.

The operator $\mathcal{D}T : \mathbf{T_X} \to \mathbf{T_X}$ sends the current tangent to the next period tangent under T. A system displays *hyperbolic dynamics*, if its tangent bundle $\mathbf{T_X}$ can be divided into two $\mathcal{D}T$ invariant sub-bundles the stable set E^S and the unstable set E^U such that the restriction $\mathcal{D}T|_{E^S}$ is a *contraction* and $\mathcal{D}T|_{E^U}$ an *expansion*. These properties are formalized as follows

1. $\mathbf{T_X} = E^S \bigoplus E^U$
2. $\mathcal{D}TE^S_X = E^S_T$ and $\mathcal{D}TE^S_X = E^S_T$ for all $X_0 \in X$
3. $\|\mathcal{D}T^r X_0\| < \rho^r \|X_0\|$ for all $v_0 \in E^S$ and $r > 0$
4. $\|\mathcal{D}T^{-r} X_0\| < \rho^r \|X_0\|$ for all $v_0 \in E^U$ and $r > 0$

where $\|\cdot\|$ is the norm formed from a suitable Riemann metric[3] and $\rho \in (0, 1)$ a contraction modulus. A similar decomposition this time for the state space surrounding a fixed point will prove useful.

Definition 22. The **stable set** W^S consists of all points whose trajectories converge in the infinite future to the steady state X^* formally

$$W^S(T, X^*) = \{X_0 \in X : \lim_{r\to\infty} T^r(X_0) = X^*\}$$

Likewise the **unstable set** W^U consists of all points whose trajectories converge in the infinite past to the steady state X^* so

$$W^U(T, X^*) = \{X_0 \in X : \lim_{r\to\infty} T^{-r}(X_0) = X^*\}$$

Note that in the second case the limit could be a set as we have not specified T be a homeomorphism.

Definition 23. In discrete time a dynamical system $T : \mathbb{R}^k \to \mathbb{R}^k$linearized system is called **hyperbolic** if it possesses no eigenvalues on the unit circle.

Remember log-linearization is isomorphic to linearization as it is a linear approximation for $X - \bar{X}/\bar{X}$ which is homeomorphic to X The following theorem indicates the power of hyperbolic dynamics for stability analysis.

[3]A Riemann metric is one that is a smooth function of both arguments of the metric- recall that this is possible because I have assumed that $\mathcal{Z}$ forms a smooth manifold.

Theorem 13. *For a hyperbolic fixed point of a smooth map $T : \mathbb{R}^k : \mathbb{R}^k$ the stable space $W^s(X_0)$ is a smooth manifold and its tangent space is the same dimension as the stable space of the linearization of T at X_0, $E^S(X_0)$.Likewise, its unstable space $W^U(X_0)$ is a smooth manifold that has the same dimension as the unstable space of the linearization of T at X_0, $E^U(X_0)$ so that $W^S \bigoplus W^U = X$.*

Proof is given in Teschl [2012] part 2.9.2 p259. [4] The next result a version of the celebrated Grobman-Hartman theorem provides the crucial application of hyperbolic dynamics to justify the use of linear approximations.

Theorem 14. *For a C^1 map $T : X \to X$ with derivative dT that has a hyperbolic fixed point at the origin there exists a homeomorphism $h(X) : N(X) \to Q(X)$ where $N(X)$ is a neighborhood of the origin for $T(X)$ and $Q(X)$ a neighborhood of the origin in the tangent space $\boldsymbol{T}_X\mathbb{R}^k$ such that $T(X) = h^{-1}(X) \circ (dT(0)) \circ h(X)$*

For proof follow p 262-266 part 2.9.3 Teschl [2012]. Observant readers will have noted that this theorem cannot be applied here because $\Delta > 1$ the following remarks treat this issue.

Remark 5. *This result naturally extends to maps defined on open intervals since these are homeomorphic (topologically equivalent) to $\mathbb{R}^k$. However, this is not the case for closed or half open intervals such as $[1, \infty)$ the state space for Δ or indeed any other intervals containing the non-stochastic steady state $\Delta = 1$. Consult corollary 4.4 to theorem 4.3 in Coayla-Teran et al. [2007] for a more abstract extension.*

The problem is that the dynamics of the benchmark New Keynesian model linearized at its non-stochastic steady state are not hyperbolic.

Proposition 14. *The log-linearized approximation exposited in Section 2.5 to the Calvo model explained in Section 2.1-2.4 does not display hyperbolic dynamics on R^k about the non-stochastic steady state.*

Proof. Recall that $X_t = (\pi_t, y_t^e)$ and the relevant operator is $T = \mathbb{E}_t V_{t+1}$. Since $\mathbb{E}_t X_{t+1} = 0$ we know that the stable linear manifold $E_T^S(0)$ of the approximate linear

[4] Since I am working with smooth functions here I have assumed smoothness- Teschl [2012] does not he prefers to work with C^q functions that are continuously differentiable k times- his proof carries over mutatis mutandis to the family of smooth functions C^∞.

system is the whole space so it is two dimensional. However, we also know that the economy has previously been in non-stochastic steady state so $X_{t-s} = 0$ for all $s > 0$ and the unstable manifold $E^U_T(0)$ must also take up the whole state space. Hence, the dimensionalities of the two linear manifolds sum to four. Formally, $\dim(E^S \bigoplus E^U) = 2 + 2 = 4 > \dim(X) = \dim(W^S \bigoplus W^U) = 2$ a contradiction of stable manifold theorem 13. $\square$

5.2.1 Valid Approximation: Stochastic Grobman Hartman

This subsection presents versions of the Grobman Hartman theorem applicable to hyperbolic fixed points in stochastic settings. There are extensions to trajectories which will be used when deriving the Phillips curve and settings that could incorporate approximate numerical solutions. The theorems come originally from Coayla-Teran et al. [2007] and they have been incorporated into Arnold [2013]. Finally, a simple argument involving topological conjugacy carries allows one to move smoothly between linearzation and log-linearization (linearization in percentage deviations.)

Theorem 15. *Let μ^* be the stochastic equilibrium of a suitably smooth cocycle $\mathbb{E}Z_{t+1} = f(Z_t, \gamma, \mathcal{U}^S_t)$ such that $N(Z)$ is a neighborhood of hyperbolic point $(Z^0, (\mathcal{U}^S)^0$ then μ^* a.e. there exists a homeomorphism such $h(Z) : N(Z) \to K(Z)$ and $W(Z)$ is a neighborhood of the tangent bundle of the origin in T^0_Z such that:*

$$f(Z, \mathcal{U}^S) = h^{-1}(f(Z^0, (\mathcal{U}^S)^0)) \circ D_{Z^0,(\mathcal{U}^S)^0)} f(\cdot)) \circ h(Z^0, (\mathcal{U}^S)^0))(Z)$$

for all Z in the domain of the composition.

This theorem is adapted from theorem one of Coayla-Teran et al. [2007]. The alterations I made were to move from a tangent space in $\mathbb{R}^n$ to its analog for a Euclidean manifold a tangent bundle[5] to reflect the fact that our ambient space is not $\mathbb{R}^n$ recall that π has poles. Also note that hyperbolicity is a property of the whole system including the error terms. As we are working with interior points here the system will be hyperbolic if and only if there are no eigenvalues on the unit circle. Its interpretation is as follows we can take linear approximations around any point we desire $(Z, \mathcal{U}^S_t)$ provided we are prepared to assume or verify hyperbolicity and be

[5] The appendix to Arnold [2013] gives a presentation of tangent bundles adequate for our purposes.

confident of getting a qualitative description of local dynamics. This holds whether or not we are at an equilibrium point.

I will take approximation only about the stochastic steady state. For two reasons firstly to facilitate comparison with perturbations about the non-stochastic steady state. Secondly so that the deviations can be interpreted as states of the business cycle. It would be possible to conceive of alternative linearization schemes that used different points of approximation. This could provide a powerful link between DSGE and the econometric literature on regime switching and asymmetric VAR. Prominent examples include Auerbach and Gorodnichenko [2012] and Tenreyro and Thwaites [2016] that address the possibilities that the effectiveness of fiscal and monetary policy vary over the business cycle, whilst Cameron et al. [1997] consider whether macroeconomic responses to post-war oil price shocks were asymmetric. It might also provide a justification for numerical methods involving linear approximation taken away from steady state such as Boppart et al. [2017].[6] Finally the theorem has powerful implications for the structure and dynamic properties of the Phillips curve.

Remark 6. *Theorem 16 implies that linearization around the stochastic steady state or any other point with hyperbolic dynamics accords with the dimensionality of the non-linear model established in Chapter 2 by Proposition 1.*

A Phillips curve approximated correctly will not look like the ZINSS Phillips curve but instead its dynamics will resemble that of the trend inflation model- even if there is no trend inflation. Conversely, the qualitative dynamics of the Rotemberg Phillips curve derived at ZINSS will accord with the true non-linear model because there is no bifurcation at this point and therefore its non-stochastic steady state is hyperbolic. The final result concerns trajectories.

Theorem 16. *For $T(Z^0, (\mathcal{U}^J)^0)$ be a hyperbolic stationary trajectory for a suitably smooth cocycle $\mathbb{E}Z_{t+1} = f(Z_t, \gamma, \mathcal{U}_t^S)$ there exists a homeomorphism $h((\mathcal{U}^J)^0) : N((\mathcal{U}^J)^0) \rightarrow W((\mathcal{U}^J)^0)$ $N(Z)$ is a neighborhood of hyperbolic point $(Z^0, (\mathcal{U}^S)^0) \rightarrow W((\mathcal{U}^J)^0)$ is a neighborhood of the origin in tangent bundle T_{Z^*} such that $T(Z^0, (\mathcal{U}^J)^0) = h^{-1}(\mathcal{U}_0^J, Z_0^J) \circ D_{Z^*} f \circ h(\mathcal{U}_0^J))(Z)$ for all Z in the domain of the composition,*

[6] Fatehi Nia and Rezaei [2017] provide a Grobman-Hartman theorem appropriate for iterated functions but it is only valid in $\mathbb{R}$ which is to restrictive for economic application.

It is an adaption of Theorem 4 in Coayla-Teran et al. [2007] with a specific random homeomorphism of interest (as defined in the paper.) This result only applies to impulse response functions- we have to take the linearization from the steady state.Its main role is to validate the application of the inverse powers of the lag operator (arbitrary forward iteration of the linear approximation) in the subsequent Phillips curve derivation. For practical purposes this result justifies forecasting from the linear model.

The final step is to verify that these results carry over to log-linearization or linearization in percentage deviations. To do so I prove the linear and log-linear solutions are topologically conjugate. As topological conjugacy is an equivalence relation this implies conjugacy between non-linear and log-linear models consistent with valid approximation.

Proposition 15. *The log-linear and linear approximations are topologically conjugate,*

Proof. Consider the demeaning transform $h : Z \to (Z_1/Z_1^*, \cdots, Z_k/Z_k^*)'$ and the linearization transform by g so $\mathbb{E}Z_{t+1} - Z^* = g(Z_t - Z^*, U_t^S)$ then by basic algebra log-linearization is represented by the transformation $h \circ g \circ h^{-1}$ which is topologically conjugate to g since $h^{-1} \circ h \circ g \circ h^{-1} = g \circ h^{-1}$ □

The intuition is that linearization and log-linearization are isomorphic because we can move from one to the other by simply adjusting coefficients leaving the dynamics of the system unchanged. The analysis for this chapter is now complete we are now free to log-linearize the New Keynesian Phillips curve about its stochastic steady state confident in its qualitative properties.

5.3 Log-Linearization and NKPC

This chapter performs log-linearization around the stochastic steady state and in particular derives an expression for the New Keynesian Phillips Curve. The crucial difference between stochastic and non-stochastic steady state approximations is that in the stochastic case the coefficients of the linear approximation will depend in general on the higher moments of the cocycle evaluated at the ergodic measure. Hence, linear approximations are not certainty equivalent as covariance and higher moments are reflected in the slope coefficients. For those well versed in non-stochastic approximations this may seem deeply counterintuitive in fact it is perhaps the most powerful aspect of stochastic approximation. It provides motivation for stochastic Grobman-Hartman theorem of the previous chapter. Finally, the dynamic properties of the stochastic model are then analyzed and compared to non-stochastic alternatives.

5.3.1 Basic Phillips Curve Solution

This subsection applies the results from the previous two chapters to log-linearize the model local to its stochastic steady state (π, π, Δ, y). It is most convenient to start with (2.35) and (2.36)

$$\begin{aligned}\hat{\aleph}_t = \eta \frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{\Delta}_t + (1+\eta)\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{y}_t - (1+\eta)\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{a}_t \\ + \frac{\alpha\beta\theta}{\aleph}\mathbb{E}(1+\pi)^{\theta-1}\aleph\,\mathbb{E}_t\hat{\pi}_{t+1} + \alpha\beta\mathbb{E}(1+\pi)^{\theta}\,\mathbb{E}_t\hat{\aleph}_{t+1}\end{aligned} \tag{5.1}$$

$$\begin{aligned}\hat{\beth}_t = \frac{\psi u'(Y)(1-\sigma)}{\beth}\hat{y}_t + \frac{\alpha\beta(\theta-1)}{\beth}\mathbb{E}(1+\pi)^{\theta-2}\beth\,\mathbb{E}_t\hat{\pi}_{t+1} \\ + \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}_t\hat{\beth}_{t+1} + \frac{\psi u'(Y)Y}{\beth}\hat{\psi}_t\end{aligned} \tag{5.2}$$

The comparative statics are instructive $\hat{\aleph}$ the business cycle deviation of weighted marginal costs is increasing in $\hat{\Delta}$ and $\hat{y}$ as these place upward pressure on real wages, whilst greater technical efficiency $\hat{a}_t$ has the opposite effect. The state of expected inflation $\mathbb{E}_t\hat{\pi}_{t+1}$ influences weighted marginal costs through the level of future sales and the next periods expectation $\mathbb{E}_t\hat{\aleph}_{t+1}$ reflects the recursive evolution of the supply side of the economy under Calvo pricing. Conversely $\hat{\beth}$ gives the deviation of the weights for the resetters marginal revenue. There is no role for price dispersion or technology here. There are forward looking terms similar to the supply side. Crucially though the effect of changes in output on marginal revenues depends on the propensity for intertemporal substitution reflected by σ (the inverse of the elasticity of substitution) if $\sigma > 1$ the propensity to smooth consumption dominates the incentive to substitution over time such that higher output today implies higher expected future marginal revenues- conversely if $\sigma < 1$ the substitution exceeds the smoothing incentive so higher current output is associated with lower future marginal revenues. When $\sigma = 1$ the two forces balance out.

To solve out the model I use the lag and expectation operators to condense the expressions for $\hat{\aleph}$ and $\hat{\beth}$ to

$$\begin{aligned}(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}\,\mathbb{L}^{-1})\hat{\aleph}_t &= \eta\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{\Delta}_t + (1+\eta)\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{y}_t \\ &- (1+\eta)\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{a}_t + \alpha\beta\theta\frac{\mathbb{E}(1+\pi)^{\theta-1}\,\aleph}{\aleph}\mathbb{L}^{-1}\hat{\pi}_t + \tilde{w}_0\mathbb{L}^{-1}\hat{u}_t^j \end{aligned} \quad (5.3)$$

$$\begin{aligned}(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\,\mathbb{L}^{-1})\hat{\beth}_t &= \frac{\psi u'(Y)(1-\sigma)}{\beth}\hat{y}_t \\ &+ \alpha\beta(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta-2}\,\beth}{\beth}\mathbb{L}^{-1}\hat{\pi}_t + \tilde{w}_1\mathbb{L}^{-1}\hat{u}_t^j \end{aligned} \quad (5.4)$$

As previously, the error terms above reflects the difference between the actual and expected values of π, $\aleph$ and $\beth$ in period $t+1$ that comes about because the future value of the structural jump errors are unknown and the model is linear in percentage deviation form. Now using the reset price equation to remove $\aleph$ and $\beth$ the price level construction equation to express the reset price in terms of inflation and then manipulating terms in the lag operator yields

$$\begin{aligned}(\mathbb{L}-\alpha\beta\mathbb{E}(1+\pi)^{\theta})(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{L}^{-1})\hat{\pi}_t &= \frac{(1-\alpha(1+\pi)^{\theta-1})}{\alpha(1+\pi)^{\theta-2}} \\ \times\Bigg[(\mathbb{L}-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})\Bigg(&\eta\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{\Delta}_t + (1+\eta)\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{y}_t \\ &- (1+\eta)\frac{\nu'(\Delta Y/A)Y}{A\aleph}\hat{a}_t + \alpha\beta\theta\frac{\mathbb{E}(1+\pi)^{\theta-1}\aleph}{\aleph}\mathbb{L}^{-1}\hat{\pi}_t\Bigg) \\ - (\mathbb{L}-\alpha\beta\mathbb{E}(1+\pi)^{\theta})\Bigg(&\frac{\psi u'(Y)(1-\sigma)}{\beth}\hat{y}_t + \alpha\beta(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta-2}\beth}{\beth}\mathbb{L}^{-1}\hat{\pi}_t\Bigg)\Bigg] \\ &+ \left(\tilde{w}_2\mathbb{L} + \tilde{w}_3 + \tilde{w}_4\mathbb{L}^{-1}\right)\hat{u}_t^j \end{aligned} \quad (5.5)$$

Expanding the lag operator, collecting terms and passing expectations from time t yields

$$
\begin{aligned}
&\hat{\pi}_{t-1} - \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}(2+\pi)\hat{\pi}_t + (\alpha\beta)^2\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta}\mathbb{E}_t\hat{\pi}_{t+1} = \\
&\quad \frac{1-\alpha(1+\pi)^{\theta-1}}{\alpha(1+\pi)^{\theta-2}}\frac{\eta\nu'(\Delta Y/A)Y}{A\aleph}\hat{\Delta}_{t-1} \\
&\quad + \frac{1-\alpha(1+\pi)^{\theta-1}}{\alpha(1+\pi)^{\theta-2}}\frac{(1+\eta)\nu'(\Delta Y/A)Y}{A\aleph}\hat{y}_{t-1} \\
&\quad - \frac{1-\alpha(1+\pi)^{\theta-1}}{\alpha(1+\pi)^{\theta-2}}\frac{(1+\eta)\nu'(\Delta Y/A)Y}{A\aleph}\hat{a}_{t-1} \\
&\quad + \beta\theta(1-\alpha(1+\pi)^{\theta-1})\frac{\mathbb{E}(1+\pi)^{\theta-1}\aleph}{(1+\pi)^{\theta-2}\aleph}\hat{\pi}_t \\
&\quad - \beta(1-\alpha(1+\pi)^{\theta-1})\frac{\mathbb{E}(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}}\frac{\eta\nu'(\Delta Y/A)Y}{A\aleph}\hat{\Delta}_t \\
&\quad - \beta(1-\alpha(1+\pi)^{\theta-1})\frac{\mathbb{E}(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}}\frac{(1+\eta)\nu'(\Delta Y/A)Y}{A\aleph}\hat{y}_t \\
&\quad + \beta(1-\alpha(1+\pi)^{\theta-1})\frac{\mathbb{E}(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}}\frac{(1+\eta)\nu'(\Delta Y/A)Y}{A\aleph}\hat{a}_t \\
&\quad - \alpha\beta^2\theta(1-\alpha(1+\pi)^{\theta-1})\mathbb{E}(1+\pi)^{\theta-1}\frac{\mathbb{E}(1+\pi)^{\theta-1}\aleph}{(1+\pi)^{\theta-2}\aleph}\mathbb{E}_t\hat{\pi}_{t+1} \\
&\quad - \frac{1-\alpha(1+\pi)^{\theta-1}}{\alpha(1+\pi)^{\theta-2}}\frac{(1-\sigma)\psi u'(Y)}{\beth}\hat{y}_{t-1} \\
&\quad - \beta(\theta-1)(1-\alpha(1+\pi)^{\theta-1})\frac{\mathbb{E}(1+\pi)^{\theta-2}\beth}{(1+\pi)^{\theta-2}\beth}\hat{\pi}_t \\
&\quad + \beta(1-\sigma)(1-\alpha(1+\pi)^{\theta-1})\frac{\psi u'(Y)}{\beth}\frac{\mathbb{E}(1+\pi)^{\theta}}{(1+\pi)^{\theta-2}}\hat{y}_t \\
&\quad + \alpha\beta^2(\theta-1)(1-\alpha(1+\pi)^{\theta-1})\mathbb{E}(1+\pi)^{\theta}\frac{\mathbb{E}(1+\pi)^{\theta-2}\beth}{(1+\pi)^{\theta-2}\beth}\mathbb{E}_t\hat{\pi}_{t+1} + \tilde{w}_5\hat{u}^j_{t-1} + \tilde{w}_6\hat{u}^j_t
\end{aligned}
\tag{5.6}
$$

removing terms in $\aleph$ and $\beth$ inside via substitution and several applications of Tonelli's theorem then simplifying coefficients and compressing errors yields

$$
\begin{aligned}
&\beta\Bigg(\alpha\mathbb{E}(1+\pi)^{\theta-1}(2+\pi)+\frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times\\
&\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}\\
&-(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y+\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Bigg)\hat{\pi}_t=\\
&\hat{\pi}_{t-1}-\frac{\eta\nu'(\Delta Y/A)Y(1-\alpha(1+\pi)^{\theta-1})/\alpha(1+\pi)^{\theta-2}A}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}\hat{\Delta}_{t-1}\\
&-\frac{(1-\alpha(1+\pi)^{\theta-1})/(1+\pi)^{\theta-2}}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}\times\\
&\Big[(1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}-(1-\sigma)\psi u'(Y)\frac{(\theta-1)}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\Big]\hat{y}_{t-1}\\
&+\frac{\beta(1-\alpha(1+\pi)^{\theta-1})/(1+\pi)^{\theta-2}\mathbb{E}(1+\pi)^{\theta-1}\eta\nu'(\Delta Y/A)Y/A}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}\hat{\Delta}_t\\
&+\frac{\beta(1-\alpha(1+\pi)^{\theta-1})/(1+\pi)^{\theta-2}}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}\times\\
&\Big[(1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1}\\
&-(1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta}\Big]\hat{y}_t\\
&+\alpha\beta^2\Bigg(\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta}+\frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times\\
&\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}\\
&-(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y+\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Bigg)\mathbb{E}_t\hat{\pi}_{t+1}\\
&+\tilde{w}_6u^j_{t-1}+\tilde{w}_7u^j_t
\end{aligned}
\tag{5.7}
$$

The expression can be compressed further by removing the two lagged terms $\hat{\Delta}_{t-1}$ and $\hat{y}_{t-1}$. To do so it is necessary to log-linearize the Euler and policy rules.

$$
\begin{aligned}
\hat{\psi}_t-\sigma\hat{y}_t=\beta&\frac{\mathbb{E}\psi u'(Y)/(1+\pi)}{\psi u'(Y)}(a_\pi\hat{\pi}_t+a_y\hat{y}_t)\\
&-\frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)}\mathbb{E}_t\hat{\pi}_{t+1}-\frac{u'(Y)}{u''(Y)}\frac{\mathbb{E}\psi u''(Y)/(1+\pi)}{\mathbb{E}\psi u'(Y)/(1+\pi)}\sigma\mathbb{E}_t\hat{y}_{t+1}
\end{aligned}
\tag{5.8}
$$

Lagging the relationship and including the later period error term yields

$$\begin{aligned}
\hat{y}_{t-1} = {} & \frac{u'(Y)}{u''(Y)}\frac{\mathbb{E}\psi u''(Y)/(1+\pi)}{\mathbb{E}\psi u'(Y)/(1+\pi)}\sigma \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\hat{y}_t \\
& - \frac{a_\pi\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi) \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\hat{\pi}_{t-1} \\
& + \frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)} \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\hat{\pi}_t \\
& + 1 \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\hat{\psi}_{t-1} \\
& - \frac{\psi\mathbb{E}u'(Y)/(1+\pi)}{\mathbb{E}\psi u'(Y)/(1+\pi)} \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\hat{\psi}_t \qquad (5.9)
\end{aligned}$$

Finally the log-linearized price dispersion relationship and its lagged form are respectively where I have used the expression for the stochastic steady of Δ

$$\begin{aligned}
\mathbb{E}_t\hat{\Delta}_{t+1} = {} & \alpha\theta\left[\mathbb{E}(1+\pi)^{\theta-1} - \frac{\mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}\Delta}\right]\mathbb{E}_t\hat{\pi}_{t+1} \\
& + \alpha\mathbb{E}(1+\pi)^{\theta}\hat{\Delta}_t \qquad (5.10)
\end{aligned}$$

$$\begin{aligned}
\hat{\Delta}_{t-1} = {} & \tilde{\omega}_8\hat{u}^j_{t-1} + \tilde{\omega}_9\hat{u}^j_t + \frac{1}{\alpha\mathbb{E}(1+\pi)^{\theta}}\hat{\Delta}_t \\
& - \theta\left(\frac{(1-\alpha)^{1/(\theta-1)}\Delta\mathbb{E}(1+\pi)^{\theta-1} - \mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}{(1-\alpha)^{1/(\theta-1)}\Delta\mathbb{E}(1+\pi)^{\theta}}\right)\hat{\pi}_t \qquad (5.11)
\end{aligned}$$

combining (7.7), (7.9) and (7.11) yields the following expression for the Phillips curve

$$\pi_t = b_0\hat{\pi}_{t-1} + b_1\hat{y}_t + b_2\hat{\Delta}_t + b_3\mathbb{E}_t\hat{\pi}_{t+1} + \tilde{\omega_{10}}\hat{u}^j_{t-1} + \tilde{\omega_{11}}\hat{u}^j_{t-1} \qquad (5.12)$$

to compact notation I use the substitution $b_i = \tilde{b}_i/b$ where b is the coefficient on inflation in (7.7) when the lagged terms are substituted out and the reset price relationship is used for further simplification

$$
\begin{aligned}
b = {} & \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}(2+\pi) \\
& + \frac{1-\alpha(1+\pi)^{\theta-1}}{\alpha(1+\pi)^{\theta-2}} \Big/ \left(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right) \\
& \quad \left[\alpha\beta\left\{\left(\theta\frac{\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right.\right.\right. \\
& \quad \left. - \beta\frac{(\theta-1)^2}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\frac{\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\right) \\
& \quad + \left((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1}\right. \\
& \quad \left. - (1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta}\right) \\
& \quad \left. \times \frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)} \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\right\} \\
& \quad - \eta\theta\nu'(\Delta Y/A)Y/A\left(\left[\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta-1}\right.\right. \\
& \quad \left. - \mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1-\alpha\mathbb{E}(1+\pi)^{\theta})\right] \\
& \quad \left.\left. /\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta}(1-\alpha\mathbb{E}(1+\pi)^{\theta})\right)\right]
\end{aligned} \tag{5.13}
$$

$$
\begin{aligned}
\tilde{b}_0 = {} & 1 + \frac{1-\alpha(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}} \Big/ \left(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right) \\
& \times \left\{\frac{a_\pi\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi) \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right)\right. \\
& \quad \times \left((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1}\right. \\
& \quad \left.\left. - (1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta}\right)\right\}
\end{aligned} \tag{5.14}
$$

$$
\begin{aligned}
\tilde{b}_1 = {} & \frac{1-\alpha(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}} \Big/ \left(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\right) \\
& \quad \times \left\{\beta\left[(1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1}\right.\right. \\
& \quad \left. - (1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta}\right] \\
& \quad - \frac{u'(Y)}{u''(Y)}\frac{\mathbb{E}\psi u''(Y)/(1+\pi)}{\mathbb{E}\psi u'(Y)/(1+\pi)}\sigma \Big/ \left(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\right) \\
& \left. \times \left((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A} - (1-\sigma)\psi u'(Y)\frac{(\theta-1)}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\right)\right\}
\end{aligned} \tag{5.15}
$$

$$\tilde{b}_2 = \frac{\eta(1-\alpha(1+\pi)^{\theta-1})}{\alpha^2(1+\pi)^{\theta-2}}\Big(- \frac{\nu'(\Delta Y/A)Y/A}{\mathbb{E}(1+\pi)^{\theta}} + \alpha^2\beta\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)//A\Big) \Big/ \Big(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Big) \tag{5.16}$$

$$\begin{aligned}\tilde{b}_3 = \alpha\beta^2\Big(&\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta} + \frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times \\ &\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A + \alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})} \\ &- (\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y + \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Big)\end{aligned} \tag{5.17}$$

By substituting the Phillips curve into the Euler we have

$$\hat{y}_t = c_0\hat{\pi}_{t-1} + c_1\hat{\pi}_t + c_2\hat{\Delta}_t + c_3\mathbb{E}_t\hat{y}_{t+1} + \tilde{\omega_{12}}\hat{u}^j_{t-1} + \tilde{\omega_{13}}\hat{u}^j_{t-1} \tag{5.18}$$

For each coefficient write $c_i = \tilde{c}_i/c$ where

$$\begin{aligned}c = \alpha\beta^2\Big(&\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta} + \frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times \\ &\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A + \alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})} \\ &- (\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y + \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Big) \\ &+ \frac{1-\alpha(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}}\Big/\Big(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Big) \\ &\times\Big\{\beta\Big[(1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1} \\ &- (1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta}\Big] \\ &- \frac{u'(Y)}{u''(Y)}\frac{\mathbb{E}\psi u''(Y)/(1+\pi)}{\mathbb{E}\psi u'(Y)/(1+\pi)}\sigma\Big/\Big(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big) \\ &\times\Big((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A} - (1-\sigma)\psi u'(Y)\frac{(\theta-1)}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\Big)\Big\} \\ &\times\frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)}\Big/\Big(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big)\end{aligned} \tag{5.19}$$

$$
\begin{aligned}
\tilde{c}_0 = -\Bigg\{ 1 + \frac{1-\alpha(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}} \bigg/ \bigg(\nu'(\Delta Y/A)\frac{Y}{A} + \alpha\beta \frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}} \bigg) \\
\times \bigg[\frac{a_\pi \beta}{\psi u'(Y)} \mathbb{E}\psi u'(Y)/(1+\pi) \bigg/ \bigg(\sigma + \frac{a_y \beta}{\psi u'(Y)} \mathbb{E}\psi u'(Y)/(1+\pi) \bigg) \\
\times \bigg((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1} \\
- (1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta} \bigg) \bigg] \Bigg\} \\
\times \frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)} \bigg/ \bigg(\sigma + \frac{a_y \beta}{\psi u'(Y)} \mathbb{E}\psi u'(Y)/(1+\pi) \bigg) \qquad (5.20)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{c}_1 = -\alpha\beta^2\Big(&\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta} + \frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times \\
&\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A+\alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})} \\
&-(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y+\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Big) \\
&\times\frac{a_\pi\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big/\Big(\sigma+\frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big) \\
&+\frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)}\Big/\Big(\sigma+\frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big) \\
&\times\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}(2+\pi) \\
&+\frac{1-\alpha(1+\pi)^{\theta-1}}{\alpha(1+\pi)^{\theta-2}}\Big/\Big(\nu'(\Delta Y/A)Y/A+\alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Big) \\
&\Big[\alpha\beta\Big\{\Big(\theta\frac{\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}} \\
&-\beta\frac{(\theta-1)^2}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\frac{\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1}}\Big) \\
&+\Big((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1} \\
&-(1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta}\Big) \\
&\times\frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)}\Big/\Big(\sigma+\frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big)\Big\} \\
&-\eta\theta\nu'(\Delta Y/A)Y/A\Big(\Big[\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta-1} \\
&-\mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1-\alpha\mathbb{E}(1+\pi)^{\theta})\Big] \\
&/\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta}(1-\alpha\mathbb{E}(1+\pi)^{\theta})\Big)\Big]
\end{aligned}
\tag{5.21}
$$

$$
\begin{aligned}
\tilde{c}_2 = -\frac{\eta(1-\alpha(1+\pi)^{\theta-1})}{\alpha^2(1+\pi)^{\theta-2}}&\Big(-\frac{\nu'(\Delta Y/A)Y/A}{\mathbb{E}(1+\pi)^{\theta}}+\alpha^2\beta\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)//A\Big) \\
&\Big/\Big(\nu'(\Delta Y/A)Y/A+\alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Big) \\
&\times\frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)}\Big/\Big(\sigma+\frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big)
\end{aligned}
\tag{5.22}
$$

$$\tilde{c}_3 = \frac{\mathbb{E}\psi u'(Y)/(1+\pi)^2}{\mathbb{E}\psi u'(Y)/(1+\pi)} \Big/ \Big(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi)\Big)$$
$$\times \alpha\beta^2\Big(\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta} + \frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times$$
$$\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A + \alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}$$
$$-(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y + \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Big) \tag{5.23}$$

For the price dispersion recursion we have

$$\hat{\Delta}_t = d_0\hat{\pi}_{t-1} + d_1\hat{\pi}_t + d_2 y_t + d_3\mathbb{E}_{t+1}\hat{\Delta}_{t+1} + \tilde{\omega_{14}}\hat{u}^j_{t-1} + \tilde{\omega_{15}}\hat{u}^j_{t-1} \tag{5.24}$$

as before write $d_i = \tilde{d}_i/d$

$$d = \alpha\beta^2\Big(\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta} + \frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}}\times$$
$$\Big[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A + \alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})}$$
$$-(\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y + \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}\Big]\Big)$$
$$-\theta\Big(\Big[\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta-1}$$
$$-\mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1-\alpha\mathbb{E}(1+\pi)^{\theta})\Big]$$
$$/\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta}(1-\alpha\mathbb{E}(1+\pi)^{\theta})\Big)$$
$$\times\frac{\eta(1-\alpha(1+\pi)^{\theta-1})}{\alpha(1+\pi)^{\theta-2}}\Big(-\nu'(\Delta Y/A)Y/A + \beta\frac{\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)//A}{\mathbb{E}(1+\pi)^{\theta}}\Big)$$
$$\Big/\Big(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}}\Big) \tag{5.25}$$

$$
\begin{aligned}
\tilde{d}_0 = \bigg\{ 1 + \frac{1-\alpha(1+\pi)^{\theta-1}}{(1+\pi)^{\theta-2}} \bigg/ \bigg(\nu'(\Delta Y/A)\frac{Y}{A} + \alpha\beta \frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}} \bigg) \\
\times \bigg[\frac{a_\pi \beta}{\psi u'(Y)} \mathbb{E}\psi u'(Y)/(1+\pi) \bigg/ \bigg(\sigma + \frac{a_y\beta}{\psi u'(Y)}\mathbb{E}\psi u'(Y)/(1+\pi) \bigg) \\
\times \bigg((1+\eta)\nu'(\Delta Y/A)\frac{Y}{A}\mathbb{E}(1+\pi)^{\theta-1} \\
- (1-\sigma)\psi u'(Y)\frac{\theta-1}{\theta}\frac{(1-\alpha)^{1/(\theta-1)}}{(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}}\mathbb{E}(1+\pi)^{\theta} \bigg) \bigg] \bigg\} \\
\times \theta \bigg(\bigg[\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta-1} \\
- \mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1-\alpha\mathbb{E}(1+\pi)^{\theta}) \bigg] \\
/\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta}(1-\alpha\mathbb{E}(1+\pi)^{\theta}) \bigg) \quad (5.26)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{d}_1 = -\theta \bigg(\bigg[\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta-1} \\
- \mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1-\alpha\mathbb{E}(1+\pi)^{\theta}) \bigg] \\
/\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta}(1-\alpha\mathbb{E}(1+\pi)^{\theta}) \bigg) \quad (5.27)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{d}_2 = \frac{\eta(1-\alpha(1+\pi)^{\theta-1})}{\alpha^2(1+\pi)^{\theta-2}} \bigg(-\frac{\nu'(\Delta Y/A)Y/A}{\mathbb{E}(1+\pi)^{\theta}} + \alpha^2\beta\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)//A \bigg) \\
\bigg/ \bigg(\nu'(\Delta Y/A)Y/A + \alpha\beta\frac{\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A}{1-\alpha\beta\mathbb{E}(1+\pi)^{\theta}} \bigg) \\
\times \theta \bigg(\bigg[\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta-1} \\
- \mathbb{E}(1+\pi)^{\theta-2}(1-\alpha(1+\pi)^{\theta-1})^{1/(\theta-1)}(1-\alpha\mathbb{E}(1+\pi)^{\theta}) \bigg] \\
/\mathbb{E}(1-\alpha(1+\pi)^{\theta-1})^{\theta/(\theta-1)}\mathbb{E}(1+\pi)^{\theta}(1-\alpha\mathbb{E}(1+\pi)^{\theta}) \bigg) \quad (5.28)
\end{aligned}
$$

$$
\begin{aligned}
\tilde{d}_3 = \frac{\beta^2}{\mathbb{E}(1+\pi)^{\theta}} \bigg(\alpha\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta} + \frac{(1-\alpha(1+\pi)^{\theta-1})}{(1+\pi)^{\theta-2}} \times \\
\bigg[\frac{\theta\mathbb{E}(1+\pi)^{\theta-1}\mathbb{E}(1+\pi)^{\theta-1}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\nu'(\Delta Y/A)Y/A + \alpha\beta\mathbb{E}(1+\pi)^{\theta}\nu'(\Delta Y/A)Y/A(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta})} \\
- (\theta-1)\frac{\mathbb{E}(1+\pi)^{\theta}\mathbb{E}(1+\pi)^{\theta-2}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})}{\psi u'(Y)Y + \alpha\beta\mathbb{E}(1+\pi)^{\theta-1}\psi u'(Y)Y/(1-\alpha\beta\mathbb{E}(1+\pi)^{\theta-1})} \bigg] \bigg) \quad (5.29)
\end{aligned}
$$

substituting in yields the canonical Phillips curve form

$$\mathbb{E}_t\hat{\pi}_{t+1} = b_0'\hat{\pi}_{t-1} + b'\hat{\pi}_t + b_1'y_t + b_2'\hat{\Delta}_t + \tilde{\omega}_{16}\hat{u}_{t-1}^j + +\tilde{\omega}_{17}\hat{u}_t^j \tag{5.30}$$

whilst the aggregate demand curve from the Euler and Taylor equation takes the form

$$\mathbb{E}_t\hat{y}_{t+1} = c_0'\hat{\pi}_{t-1} + c_1'\hat{\pi}_t + c'y_t + c_2'\hat{\Delta}_t\tilde{\omega}_{18}\hat{u}_{t-1}^j + +\tilde{\omega}_{19}\hat{u}_t^j \tag{5.31}$$

using $\hat{\Delta}_t = d_0'\hat{\pi}_t/(1-d_1'\mathbb{L}) + \tilde{\omega}_{19}\hat{u}_t^j$. I can solve for a VARMA representation by lagging each relationship by a period, thus

$$\hat{y}_t = c_4'\hat{\pi}_{t-1}+c_5'\hat{\pi}_{t-2}+c_6'\pi_{t-3}+c_7'\hat{y}_{t-1}+c_8'\hat{y}_{t-2}+\tilde{\omega}_{20}\hat{u}_{t-3}^j+\tilde{\omega}_{21}\hat{u}_{t-2}^j+\tilde{\omega}_{22}\hat{u}_{t-1}^j+\tilde{\omega}_{23}\hat{u}_{t-1}^j \tag{5.32}$$

$$\hat{\pi}_t = b_4'\hat{\pi}_{t-1}+b_5'\hat{\pi}_{t-2}+b_6'\hat{\pi}_{t-3}+b_7'\hat{y}_{t-1}+b_8'\hat{y}_{t-2}+\tilde{\omega}_{24}\hat{u}_{t-3}^j+\tilde{\omega}_{25}\hat{u}_{t-2}^j+\tilde{\omega}_{26}\hat{u}_{t-1}^j+\tilde{\omega}_{27}\hat{u}_t^j \tag{5.33}$$

This is a Vector Autoregressive Moving Average- VARMA (p,q) where

$$\mathbf{X}_t = A\sum_{i=0}^{p}\mathbf{X_{t-i}} + \mathbf{B}\sum_{\mathbf{j=0}}^{\mathbf{q}}\mathbf{e_{t-j}}$$

here $\mathbf{X} = (\hat{y},\hat{\pi})$, $\mathbf{e} = (\hat{\psi},\hat{a})$ and $p = q = 3$ Further parametized coefficients reported on request.

Chapter 6

Thesis Conclusion

This thesis shows that there are problems with the approximation of the benchmark Calvo New Keynesian model at its non-stochastic steady state. This is clear because the underlying system has a canonical form with a rich persistence structure; it can be represented as a VARMA (3,3). This is not equivalent to the reduced form approximation taken at the non-stochastic steady state, which is the trivial VAR (0). This is because there are non-hyperbolic dynamics around the non-stochastic steady state. I relate this to bifurcation and the novel notion of a weak non-negativity constraint all properties local to this extreme point of the state space.

I have derived the Phillips curve. It generates persistence and will allow me to analyze policy trade-offs with the potential for substantial implications for optimal policy. I find a compelling practical motivation to substantially extend and dramatically relax the economic notion of equilibrium. It has the benefit of aligning economics with other sciences. Economics will always be dominated by optimization since it is how we model human responses to well-defined incentives. However, constant optimization is a shibboleth and it will surely prove useful for economic theory to have an implement in its toolkit to model situations where its absence is potentially significant. It may also allow greater plurality of ideas and arguments to enter the mainstream macroeconomic debate.

Furthermore, the thesis offers axiomatic treatments of coordination failure arising from price rigidity as well as the Lucas critique, that I hope further economic understanding and overturn counterintuitive results such as the Divine Coincidence. Finally, the paper outlines a sophisticated arsenal of mathematical techniques from topology and ergodic theory that are unfamiliar to most economic theorists particularly in macroeconomics. The comparative statics techniques devised in Theorem 13 and Proposition 11 could have significant application across wide areas of modern economics. Finally, I hope my research will allow fruitful collaboration between economists and mathematicians and mathematical physicists in areas hitherto seen as unrelated.

I continue to work on further details. The greatest challenge I have had with the thesis is the size of the problem to be tackled. Dynamic Stochastic General Equilibrium models are an order of magnitude larger than models studied by micro-theorists. I have shown that the mathematics involved is extremely challenging and arguably substantially more wide-ranging and advanced than the techniques currently popular in microeconomics. I believe the thesis can lead to a complete solution to the canonical New Keynesian model including plausible optimal policy. However, I feel I have given readers a flavor of the kind of techniques and arguments required to attack this problem, perhaps the largest and most challenging our profession has to offer today. Getting the New Keynesian framework right is of crucial importance to the conduct of macroeconomic policy, public understanding of the business cycle and in so far as one can devise new schemes to manage it, social welfare. Furthermore, it is of central importance to the economic profession. The scientific achievements of other branches of economics are frequently undermined in public perception by the widely-supposed failures of macroeconomists. I have shown there is significant truth to this position in terms of DSGE models used for monetary policy. This line of research that treats DSGE as a dynamical system has the potential to win back public respect for academic macroeconomics, with positive spillovers for the rest of the profession.

Chapter 7

Appendices

The appendix is organized into two sections. The first covers results from Chapter 2 the second results from Chapter 3.

7.1 Results from Chapter 2

The first two subsections cover comparative statics from Section 2.2 in particular Subsection 2.2.3. The third extends the policy analysis in 2.4. The final two relate to arguments in 2.6 **Stochastic Averaging**

7.1.1 Mathematics of GSM

As far as I understand GSM is a new mathematical concept however it is perfectly congruous with existing work on *generalized mean* which tries to search for general properties common to all measures of central tendency that depend upon all observations in the sequence. GSM encompasses other 'average concepts' such as the weighted average $\mathcal{M} = \sum_i w_i x_i$, the generalized power mean $\mathcal{M}(\rho_\prime) = \lim_{\rho \to \rho_0} [\sum x_i^\rho]^{1/\rho}$ where $\rho \in \bar{\mathbb{R}}$ where special cases include $\rho_0 = 0$ corresponds to the geometric mean $\Pi(x_i \ldots x_n)^{\frac{1}{n}}$ and the more general f mean $\mathcal{M}_f = f^{-1}(\sum_i x_i)$ which need not be homogeneous and compositions of these consult Means and Means [2003] for a broader perspective on this literature. It possesses two useful properties.

There are several points worth discussing about the definition. The first is that condition one only applies to non-stochastic sequences this is to swerve around Jensen's inequality which ensures that $\mathcal{M}\mathbb{E}[x] \neq \mathbb{E}[\mathcal{M}\mathbb{E}[x]]$ For example the generalized power mean order 2 $\sum_i (x_i^2)^{1/2}$ is a GSM on $\mathbb{R}$ but as a composition of a norm (the Euclidean) and a linear function it is a convex function so $\mathcal{M}\mathbb{E}[x] \geq \mathbb{E}[\mathcal{M}\mathbb{E}[x]]$ for proof see Boyd and Vandenberghe [2004]. The strictly increasing requirement in condition (ii) rules out the median and the mode that in general do not depend on all the variables. The continuity restriction rules out functional dependence upon the mode which is a measure of tendency but not central tendency; since with non-convexities in the sample space distribution the mean might never occur but the mode might be the minimum value. Continuity also ensures that a 'generalized mean' represents the underlying sequence with small changes in one leading to small changes in the other and vice versa. Recall that consistent with the construction of the optimization problem of an individual firm which is small relative to the aggregate economy it is the case that $\frac{\partial \mathbb{E}(f(x))}{\partial x} = \mathbb{E}f'(x)$. I assume throughout that the stochastic process admits the existence of $\mathcal{D}[\mathbb{E}f)]$ wherever its non-stochastic counterpart exits. Hence the expectations operator is passive here. However, as it is not primary interest here I will not pursue this here as the properties of averages is not my primary line of inquiry. If a linear approximation is taken then this GSM is equivalent to a more conventional weighted average of expected future reset prices, although valid only locally. This is a well known result its use in economics dating back at least to Fischer [1977]. Since I go onto distinguish between linear and non-linear models this approach is inconvenient for me.

7.1.2 Proof Lemma 1

There are two cases to consider, the first with constant returns to scale is simpler as actual marginal costs will not differ from those of a flexible price firm. The second case where there are decreasing returns to scale requires establishing a mapping between firms' marginal costs under rigid and flexible prices. This section draws heavily on concepts in infinite dimensional analysis and measure theory. For those with minimal background Aliprantis and Border [2006] Chapters 6, 7 and 10 would be a good starting point- pages 227-229, 274 and 391 provide the basic definitions. **Case 1: Constant Returns to Scale** Begin by substituting [18] into [19] which yields:

$$(p_t^*)^{-\theta}\mathbb{E}_t \sum_{T=t}^{\infty} w_T \left[(p_t^*) - \frac{\theta}{\theta-1} p_T^f\right] = 0$$

where the weight $w_T = (1-\alpha)^{T-t} Q_{t,T} Y_T (P_T)^{\theta-1}$ is determined by aggregate variables and is therefore *independent* of the price-setting problem of an individual firm. Condition [i] follows because when $p_t^f = p^f$ by construction of the price level $p_t^f = P_t$, then substituting into the first order condition of the optimal reset price problem [18] yields the flexible price problem with $p_t^* = p_t^f$.

The existence and uniqueness of the function $\mathcal{A}$ requires differentiation and an implicit function argument, whilst condition [ii] follows from inspecting partial derivatives. To do so I introduce an auxiliary variable to enable implicit differentiation

$$\mathcal{B}_t(p_t^*, \mathcal{P}_t^f; \mathcal{S}, \Omega_t) = (p_t^*)^{-\theta} \mathbb{E}_t \sum_{T=t}^{\infty} w_T [p_t^* - \frac{\theta}{\theta - 1} p_T^f]$$

where $\mathcal{P}_t^f = \langle p_t^f, p_{t+1}^f, \cdots \rangle$ is the future path of optimal flexible prices and $\mathcal{S}$ the stochastic process including but not limited to terms in w and Ω_t is all information about the process up to time t. The state spaces are $p_t^f \in \mathfrak{P}^f$, $\mathcal{P}_t^f \in \mathfrak{P}^f \times \mathfrak{P}^f \cdots$, $p_t^* \in \mathfrak{P}^*$ and $\mathcal{B}_t \in \mathfrak{B}$. As $\mathcal{B}_t$ and p_t^* are non-stochastic variables their state spaces have the Euclidean norm $\|.\|_E$ corresponding to the familiar notion of 'distance' from the origin. For the stochastic variables- p_t^f and $\mathcal{P}_t^f$ the norms are respectively the L^1 norm $\|p_t^f\| = \int_{\mathfrak{P}^f} |p_t^f| \mathrm{d}\mu[S, \Omega_t]$ and $\|\mathcal{P}_t^f\|_{\mathfrak{P}} = \mathbb{E}_t \sum_{T=t}^{\infty} w_T \|p_T^f\|$ where the expectation is taken with respect to the product measure over $\mathfrak{P}^f$ defined because the function $\mathcal{B}$ forms a Hilbert space. Now since $\mathcal{B}$ is comprised of the sum of functions continuously differentiable in the relevant variables (in section 2 I assumed MC had two derivatives and other functions are power functions which are smooth) we know that it's Fréchet derivative exists and is continuous. Recall that the Fréchet derivative $D\mathcal{B}$ is a generalization of the derivative to Banach spaces such that

$$D\mathcal{B}_t \equiv \lim_{\delta \to 0} \frac{\mathcal{B}_t(p_t^* + \delta h, \mathcal{P}_t^f + \delta h; \mathcal{S}_t, \Omega_t) - \mathcal{B}_t(p_t^*, \mathcal{P}_t^f; \mathcal{S}_t, \Omega_t)}{\delta}$$

The final condition to apply the implicit function theorem is that $\mathcal{B}_t$ and p_t^* be diffeomorphic local to zero which follows from

$$\frac{\partial \mathcal{B}_t}{\partial p_t^*}|_{\mathcal{B}_t=0} = -\theta \frac{\mathcal{B}_t}{p_t^*} + (p_t^*)^{-\theta} \mathbb{E}_t \sum_{T=t}^{\infty} w_T = (p_t^*)^{-\theta} E_t \sum_{T=t}^{\infty} w_T > 0$$

The full proof is laid out in Lang [2001] p15-21. Finally condition [ii] is verified by implicit differentiation.

$$\mathbb{E}_t \frac{\partial \mathcal{B}_t}{\partial p^f_\tau} = -\frac{\theta}{\theta - 1}(p^*_t)^{-\theta}\mathbb{E}_t w_\tau < 0$$

$$\mathbb{E}_t \frac{\partial p^*_t}{\partial p^f_\tau} = -\mathbb{E}_t[\frac{\partial \mathcal{B}_t}{\partial p^f_T}/\frac{\partial \mathcal{B}_t}{\partial p^*_t}] = \frac{\theta}{\theta - 1}\frac{\mathbb{E}_t w_\tau}{\mathbb{E}_t \sum_{T=t}^{\infty} w_T} > 0$$

Case 2: Decreasing Returns to Scale

Here marginal costs are increasing in output so in general there will be a difference between marginal costs paid at the flexible price output and sticky prices. The proof follows swiftly from establishing an increasing relationship between marginal costs of every firm whatever its price. Call this relationship $\mho$ so

$$MC_t(i)\left(\frac{p_i}{P};\Omega_t\right) = \mho\left(\frac{p_i}{P};\Omega_t\right)$$

$$MC^*_t = \mho(1;\Omega_t)$$

we know from 2.2.1 this is increasing and differentiable in the second argument (which includes output and an appropriate representation of shocks and parameters.) Hence for every real price of an individual firm i there is a diffeomorphism $H(\frac{p_i}{P}) : \Omega_t \to MC_t(i)(\frac{p_i}{P};\Omega_t)$ hence $\mho = H(\frac{p_i}{P}) \circ H^{-1}(1))$ which retains the property of being an increasing diffeomorphism hence

$$\mathbb{E}_t \frac{\partial \mathcal{B}_t}{\partial p^f_\tau} = -\frac{\theta}{\theta - 1}(p^*_t)^{-\theta}\mathbb{E}_t w_\tau H'_\tau < 0$$

$$\mathbb{E}_t \frac{\partial p^*_t}{\partial p^f_\tau} = -\mathbb{E}_t \frac{\partial \mathcal{B}_t}{\partial p^f_\tau}/\frac{\partial \mathcal{B}_t}{\partial p^*_t} = \frac{\theta}{\theta - 1}\frac{\mathbb{E}_t w_\tau H'_\tau}{\mathbb{E}_t \sum_{T=t}^{\infty} w_T} > 0 \quad \square$$

Derivation of Remark 1

Proof. $\mathcal{B}_t$ can be rewritten as

$$\mathcal{B}_t = \left(\frac{p^*_t}{P_t}\right)^{-\theta}\sum_{T=t}^{\infty}(1-\alpha)^{T-t}Q_{t,T}Y_T\Pi^{\theta-1}_{t,T}\left[\frac{p^*_t}{P_t} - \Pi_{t,T}MC_T\left(\frac{p^*_t}{P_t}\Pi_{t,T}\right)\right] > 0$$

I proceeding as before by an implicit function argument:

$$\mathbb{E}_t \frac{\partial \mathcal{B}_t}{\partial \Pi_{t,T}}|_{\mathcal{B}_t=0} = -\mathbb{E}_t\Big[(\theta-1)\frac{\mathcal{B}_t}{\Pi_{t,T}} + \Big(\frac{p_t^*}{P_t}\Big)^{-\theta} \sum_{T=t}^{\infty}\Big((1-\alpha)^{T-t} Q_{t,T} Y_T \Big(MC_T \ - \ y_T'(i)\Big)\Big(\frac{p_t^*}{P_t}\Big)\Big(\frac{1}{\Pi_{t,T}}\Big)^2 MC_T'(i)\Big)\Big] < 0$$

$$\mathbb{E}_t \frac{\partial \mathcal{B}_t}{\partial (p_t^*/P_t)}|_{\mathcal{B}_t=0} = \mathbb{E}_t\Big[-\theta\frac{\mathcal{B}_t}{p_t^*/P_t} + \Big(\frac{p_t^*}{P_t}\Big)^{-\theta}\sum_{T=t}^{\infty}(1-\alpha)^{T-t} Q_{t,T} Y_T (\Pi_{t,T})^{\theta-1}\Big(1 \ - \ y_T'(i) MC_T'(i)\Big)\Big] > 0$$

$$\mathbb{E}_t \frac{\partial (p_t^*/P_t)}{\partial \Pi_{t,T}} = \mathbb{E}_t\Big[\frac{\sum_{T=t}^{\infty}(1-\alpha)^{T-t} Q_{t,T} Y_T (MC_T \ - \ y_T'(i)(\frac{p_t^*}{P_t})(\frac{1}{\Pi_{t,T}})^2 MC_T'(i))}{\sum_{T=t}^{\infty}(1-\alpha)^{T-t} Q_{t,T} Y_T (\Pi_{t,T})^{\theta-1}(1 \ - \ y_T'(i) MC_T'(i))}\Big] > 0$$

□

Understanding Policy Rule

7.1.3 Taylor (1979)

In keeping with the style of the time Taylor simply laid out log-linear equations without an underlying non-linear model- in keeping with Chapter 2 I consider price rather than wage contracts and simplified some familiar features such as discounting. The first equation states that the deviation of the reset price $\hat{p}_t^*$ from (non-stochastic) equilibrium is set as a weighted average of expected future realizations of the deviation of the optimal flexible price deviation $\hat{p}_t$ throughout the (M period) duration of the contract.

$$\hat{p}_t^* = \frac{1}{M}(\hat{p}_t + \cdots + \mathbb{E}_t \hat{p}_{t+M-1})$$

The price level is equivalent to the optimal flexible price deviation as firms problems are identical hence

$$\hat{p}_t = \frac{1}{M}(\hat{p}_t^* + \cdots + \mathbb{E}_t \hat{p}_{t-(M-1)}^*)$$

To solve out for the inflation gap $\hat{\pi}_t = \hat{p}_t - \hat{p}_{t-1}$ substitute out the optimal reset price using the lag operator L where $L^m x_t = x_{t-m}$ and $L^{-m} x_t = E_t x_{t+m}$ for any $m \geq 0$

on the first equation yields

$$\hat{p}_t^* = \frac{1}{M}\frac{(1-L^{-M})}{(1-L^{-1})}\hat{p}_t$$

Now note that $1 - L^M = 1 - L + L - L^2 + L^2 - L^3 + \cdots + L^{M-1} - L^M$ implies

$$(1 - L^M)\hat{p}_t = \pi_{t-1} + \pi_{t-2} + \cdots + \pi_{t-(M-1)}$$

$$(1 - L^{-M})\hat{p}_t = \pi_{t+1} - \pi_{t+2} - \cdots - \pi_{t+M}$$

which yields the final result

$$\mathbb{E}_t\pi_{t+1} = \pi_t + \frac{1}{m^2-1}\underbrace{\mathbb{E}_t[(\pi_{t+2} - \pi_{t-1}) + (\pi_{t+q} - \pi_{t-(q+1)}) + \cdots + (\pi_{t+M} - \pi_{t-(M-1)})]}_{\text{q+1 < M}}$$

The expectation of next period inflation $E_t\pi_{t+1}$ is a function of past and future and inflation and the canonical expression below confirms that it is different from zero in general and can therefore give rise to provide the basis for a non-trivial forward-looking monetary policy.

$$\pi_{t+1} = -\pi_t - \cdots - \pi_{t-(M-3)} + (m^2-1)(\pi_{t-(M-2)} - \pi_{t-(M-1)}) + \pi_{t-M} + \cdots + \pi_{t-(2M-2)}$$

By altering the time scale to annual frequency the expectation structure advocated by Taylor in defense of his rule would be correct if contracts lasted two years (eight quarters) and the annual inflation rate $\hat{\pi}_t^* = \hat{p}_t - \hat{p}_{t-4}$, $\mathbb{E}_t\hat{\pi}_{t+1}^* = \hat{p}_{t+4} - \hat{p}_t$ were the inflation measure as then

$$\mathbb{E}_t\pi_{t+1}^* = \pi_t^*$$

as Taylor argued. In reality the average contract length tends to be shorter around three to four quarters in European nations and two to three quarters in the United States compare Dhyne et al. [2006] and Alvarez et al. [2006] with Bils et al. [2004][1].

[1] Naturally, if we believed US firms changed their prices every two quarters and we altered the inflation measure to the quarterly rate as standard in our models then the unit root result would retain and Taylor's expectation structure would be correct. With wage rigidity instead or in combination with price rigidity the dynamics of these model for given contract length are analogous up to the interpretation of coefficients- see Edge [2002] orAscari [2003]. Dixon and Le Bihan [2012] show that the weighted average length of wage spells in a representative French Panel dataset is two quarters. Although this maybe anomalous for my purposes since Dickens et al. [2007] using annualized data shows that nominal wages in France are substantially more flexible than those in most other OECD nations, including in particular the United States.

Naturally, there are methodological doubts about the practice of calibrating stylized models with micro data. This is why Taylor's rule was always marketed as a guide to policy rather than a strictly optimal policy which depends on model specific details. A literature on robustness of rules across models and optimal simple rules follows this lead. Rudebusch [2001], Onatski and Stock [2002], Levin et al. [2003], Levin et al. [2006], Žaković et al. [2007], Brock et al. [2007] and Sala et al. [2008] are examples of major papers on robustness, whilst for optimal simple rules consult Schmitt-Grohe et al. [2007],Kumhof et al. [2010], Taylor et al. [2010] and Orphanidesa and Wielandb [2013]. The two are frequently linked with simple rules tending to be more robust to model misspecification and parameter uncertainty Levin et al. [1999], Levin et al. [2007] and Taylor et al. [2010] detail this argument. **Results from Section 2.6** This subsection presents more details of the derivations related to the persistence problem with the benchmark New Keynesian model, alongside several extensions which indicate the robustness of the puzzle.

7.1.4 Eigenvalues and Convergence

To begin with the expression for the eigenvalues of the matrix A is as follows.

$$\lambda_1 = \frac{\sigma^{-1}(a_y + \omega\beta^{-1}) + 1 + \beta^{-1} - \sqrt{[\sigma^{-1}(a_y + \omega\beta^{-1}) + 1 + \beta^{-1}]^2 - 4\beta^{-1}(1 + \sigma^{-1}(a_y + \omega a_\pi))}}{2}$$

$$\lambda_2 = \frac{\sigma^{-1}(a_y + \omega\beta^{-1}) + 1 + \beta^{-1} + \sqrt{[\sigma^{-1}(a_y + \omega\beta^{-1}) + 1 + \beta^{-1}]^2 - 4\beta^{-1}(1 + \sigma^{-1}(a_y + \omega a_\pi))}}{2}$$

Note that both are positive and the larger eigenvalue λ_2 is always greater than one. We need both to be outside the unit circle for unconditional convergence. In the case where the discriminant term under the square root is negative the solution takes the form $x_t = e^{-\gamma t}(Acos(zt) + Bsin(zt))$ which will converge non-monotonically when the real part of both eigenvalues is outside the unit circle ($\gamma > 0$) see for example Simon and Blume [1994] for an exposition of this case.

7.1.5 Full Solution for Benchmark Model

The coefficients for inflation are

$$\zeta_\pi^1 = \frac{-\sigma\beta\omega^2(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)})}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

$$\zeta_\pi^2 = \frac{\beta\omega^2(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)})}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

$$\zeta_\pi^3 = \frac{\sigma\omega[(\beta\lambda_2 - 1)\lambda_1^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_2^{-(1+i)}]}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

For output we have

$$\zeta_y^1 = \frac{\sigma\beta\omega[(\beta\lambda_2 - 1)\lambda_2^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_1^{-(1+i)}]}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

$$\zeta_y^2 = \frac{-\beta\omega[(\beta\lambda_2 - 1)\lambda_2^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_1^{-(1+i)}]}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

$$\zeta_y^3 = \frac{\sigma(\beta\lambda_1 - 1)(\beta\lambda_2 - 1)(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)}) - -\omega[(\beta\lambda_2 - 1)\lambda_2^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_1^{-(1+i)}]}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

Finally for interest rates

$$\zeta_i^1 = \frac{-a_\pi\sigma\beta\omega^2(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)}) + a_y\sigma\beta\omega[(\beta\lambda_2 - 1)\lambda_2^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_1^{-(1+i)}]}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

$$\zeta_i^2 = \frac{a_\pi\beta\omega^2(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)}) - a_y\beta\omega[(\beta\lambda_2 - 1)\lambda_1^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_2^{-(1+i)}]}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

$$\zeta_i^3 = \frac{\begin{array}{l} a_\pi\beta\omega[(\beta\lambda_2 - 1)\lambda_1^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_2^{-(1+i)}] + a_\pi\omega^2(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)}) \\ \quad + a_y\sigma(\beta\lambda_1 - 1)(\beta\lambda_2 - 1)(\lambda_2^{-(1+i)} - \lambda_1^{-(1+i)}) - a_y\omega(\beta\lambda_2 - 1)\lambda_2^{-(1+i)} - (\beta\lambda_1 - 1)\lambda_1^{-(1+i)} \end{array}}{\sigma\beta^2(\lambda_2 - \lambda_1)}$$

None of the three interest rate coefficients can be decisively signed which explains the counter-intuitive or uncertain signs of all other coefficients.This confirms the interpretation in the text that the error terms do not correspond with our intuition of what a shock is when policy responds to contemporaneous variables. For example, $\zeta_\pi^1 < 0$ implies that a preference shock which causes the household to move consumption from the next period to the present actually causes present consumption to *fall*. This is a consequence of the increase in real interest rate associated with the Taylor principle. When there are repeat eigenvalues the form of the solution is different.

The repeat eigenvalue is $\lambda = \sigma^{-1}(a_y + \omega\beta^{-1}) + 1 + \beta^{-1}$ as many terms are zero in the expansion we find that see Halmos [1958]

$$A^{-(1+i)} = -i\lambda^{-(1+i)} + (1+i)\lambda^{-i}A^{-1}$$

. General solution equation (37) then allows us to calculate the ζ coefficients.

7.2 Proofs from Chapter 3

This item proves the three propositions about price dispersion in the stochastic New Keynesian model mentioned in Chapter 3.

7.2.1 Proof of Proposition 5

Proof. The proof that $\Delta \geq 1$ is an application of Jensen's inequality. First define two functions.

$$g(p_i) = p_i * \left(\frac{p_i}{P}\right)^{-\theta}$$

$$\phi(p_i) = \left(\frac{p_i}{P}\right)^{\theta/(\theta-1)}$$

We need to assign a probability measure for the prices. Any non-singular measure assigning zero probability $p_i = 0$ at every history will suffice. Note that the construction of the price level ensures that then $P > 0$ with probability one. Therefore, we know that ϕ is strictly convex on every measurable set since $\frac{d^2\phi}{dp_i{}^2} = \frac{\theta}{(\theta-1)^2}\frac{\phi(p_i)}{p_i^2} > 0, \forall\, p_i > 0$. Although in the first part of the proof I will only use the weak convexity property. Note that $P = \int_\Omega g \; \mathrm{d}\mu$. Now since ϕ is a convex function defined on a metric space it follows from theorem 7.12 (p. 265) in Aliprantis and Kim that it has a sub-derivative at every point. Hence we may select a and b such that:

$$ap^* + b \leq \phi(p^*)$$

For all possible reset prices $p*$ and for the particular value $p^* = P$

$$aP + b = \phi(P)$$

It follows that:

$$\phi \circ g(p^*) \geq ag(p^*) + b$$

For all p^* since we have a probability measure the integral is monotone with $\mu(\Omega) = 1$.
Note that:

$$\begin{aligned}
\Delta &= \int_\Omega \phi \circ g \, \mathrm{d}\mu \\
&\geq \int_\Omega (ag + b) \, \mathrm{d}\mu \\
&= a \int_\Omega g \, \mathrm{d}\mu + \int_\Omega b \, \mathrm{d}\mu \\
&= aP + b \\
&= \phi(P) \\
&= 1
\end{aligned}$$

□

Where I have used respectively the monotonicity of the Lebesgue integral, the linearity of the Lebesgue integral and the definition of the functions see Royden and Fitzpatrick (p80-82).

For the second part we need to be clear about the nature of the measure used. It is a discrete measure corresponding to price dispersion statistics defined at each time t. As a probability measure it is defined by the triplet (Ω, Σ, μ) where Ω is a probability space Σ is a sigma-algebra of sets and μ is a probability measure defined for every set in Σ. Ω_t is the set of all prices in the economy at time t. Σ_t is the set of all subsets (or power set) of Ω_t denoted $\Sigma_t = \mathcal{P}(\Omega_t)$ and μ_t is the share of particular prices in the economy at time t. To generalize the result across models simply modify the probability measure. The definition of Σ_t will be the same for all discrete time models but Ω_t and μ_t will change. For the Calvo model without indexation used here they are as follows:

$$\Omega_t = \{\cdots, p^*_{-1}, p^*_0, p^*_1, \cdots, p^*_t\}$$

$$\mu_t(p) = \Sigma^t_{-\infty} \delta_T(p) \alpha^{t-T}(1 - \alpha)$$

Where p^*_T indexes the reset price at time t and δ_T is the indicator function for time T.

Defined as $\delta_T(p) = \begin{cases} 1 & \text{if } p = p_T^* \\ 0 & \text{otherwise.} \end{cases}$ $\forall\, 0 \leq T \leq t$ Note that the incongruous feature of having an infinite history of reset prices is not necessary to prove Lemma 1 for the Calvo model - the result would pass with a finite history of reset prices starting at p_0^* with the measure $\mu_t(p) = \Sigma_{T=1}^t \delta_T(p)\alpha^{t-T}(1-\alpha) + \delta_0(p)\alpha^t$. For the Yun [1996] model where the Calvo pricing firms not allowed to re-optimize, index to trend inflation $\bar{\pi}$ they are:

$$\Omega_t = \{\cdots,\ (1+\bar{\pi})^{t+1}p_{-1}^*(1+\bar{\pi})^t p_0^*,\ (1+\bar{\pi})^{t-1}p_1^*, \cdots,\ p_t^*\}$$

$$\mu_t(p) = \Sigma_{-\infty}^t \hat{\delta}_{t,T}(p)\alpha^{t-T}(1-\alpha)$$

Where $\hat{\delta}_{t,T}(p) = \begin{cases} 1 & \text{if } p = (1+\bar{\pi})^{t-T}p_T^* \\ 0 & \text{otherwise.} \end{cases}$ $\forall\, 0 \leq T \leq t$

For the Generalized Taylor Economy:

$$\Omega_t = \{\cdots,\ p_{t-(J-1),J}^*,\ p_{t-(J-2),J}^*,\ p_{t-(J-2),J-1}^*, \cdots,\ p_{t-1,2}^*,\ p_{t-1,3}^*, \cdots,\ p_{t-1,J}^*,\ p_{t,1}^*,\ p_{t,2}^*, \cdots,\ p_{t,J}^*\}$$

$$\mu_t(p) = \sum_{j=1}^{J}\sum_{k=0}^{j-1} \delta_{t-k,j}\gamma_j/j$$

Note that the second subscript now indicates contract length and J is the maximum contract length. γ_j is the share of firms with contract length j. There is staggered pricing for each contract length so $1/j$ is the fraction of firms with contract length j resetting at a given time. The need to generate staggered nominal adjustment within each sector is the reason why we cannot begin the contract history at the date $t = 0$. Note that unlike with Calvo there is not a single reset price in each period but one for each contract length 1 through j.

Finally, other models of state dependent pricing can be admitted trivially by making parameters in the measure dependent on the parameters and history of shocks.

To derive the conditions for $\Delta > 1$ we need to employ the strict convexity property of ϕ which is as follows:

$$\phi(sp + (1-s)p') < s\phi(p) + (1-s)\phi(p')$$

for any s in (0,1) and $p \neq p'$. Select any p' we know from its weak convexity property that ϕ has a sub-derivative at p' so:

$$ap + b \leq \phi(p)$$

$$ap' + b = \phi(p')$$

Now here begins a contradiction argument. Suppose another point p'' had the same sub-derivative then:

$$ap'' + b = \phi(p'')$$

Now consider a convex combination $p''' = sp' + (1-s)p''$. As ϕ is strictly convex we know that:

$$s\phi(p') + (1-s)\phi(p'') > \phi(p''')$$

However, as we have assumed they have the same subderivative we obtain a contradiction:

$$s(ap' + b) + (1-s)(ap'' + b) = ap''' + b > \phi(p''')$$

Therefore for all $p^* \neq P$:

$$\phi \circ g(p^*) > ag(p^*) + b$$

Now partition the sample space as follows: $\Omega_t^1 = \Omega_t \backslash \{p_T^* : p_T^* = P_t\}$ and $\Omega_t^2 = \Omega_t \backslash \Omega_t^1$ This allows me to decompose the condition for $\Delta = 1$ as follows:

$$\int_{\Omega_t^1} \phi \circ g \, \mathrm{d}\mu + \int_{\Omega_t^2} \phi \circ g \, \mathrm{d}\mu = \int_{\Omega_t^1} (ag + b) \, \mathrm{d}\mu + \int_{\Omega_t^2} (ag + b) \, \mathrm{d}\mu$$

Now we can cancel the second expression on each side because we know they are always equal on every set in Ω_t^2 and we can invoke uniform continuity of the Lebesgue integral to extend this to the sigma-algebra. Rearranging and reparametizing with the function $h = \phi \circ g - (ag + b)$ yields the condition:

$$\int_{\Omega_t^1} h \, \mathrm{d}\mu = 0$$

Now the next step is to prove that the set Ω_t^1 has measure zero. Since $h \geq 0$, I can apply Chebyshev's inequality which states that for any $\epsilon > 0$:

$$\mu(\{h > 0\}) \leq \frac{1}{\epsilon} \int_{\Omega_t^1} h \, \mathrm{d}\mu = 0$$

. Take the union over sequences $\epsilon_k \searrow 0$ to obtain $\mu(\{h > 0\}) = 0$. Now note that:

$$p \in \Omega_t^1 \Rightarrow h(p) > 0 \Rightarrow p \in \{h > 0\}$$

So $\Omega_1^t \subseteq \{h > 0\}$ thus $\mu(\Omega_t^1) = 0$.Now this means every subset of Ω_t^1 must have zero probability including every individual reset price p however this contradicts the definition of Ω that a positive fraction of firms are selling at price p. Therefore Ω_t^1 is the empty set and $\Omega_t^2 = \Omega$, so $\Delta = 1$ if and only if every reset price $p^* = p$. To prove $p_i = P$ write the price level as $P = \int_i p_i(\frac{p_i}{P})\mathrm{d}\mu$.Note that with all firms setting the same price p_i is independent of μ which with a dispersed price level would reflect the share of firms resetting prices at a certain date. We can therefore factorize p_i from the integral to leave $P = p_i\Delta$ since I have already shown that $\Delta = 1$ when all $p_i = P$, the proof is complete.

7.2.2 Definition 1 and Lemma 3 Material

This subsection contains extensions of Lemma 3 to cover various nominal indexation schemes proposed in the literature and the example cited in the text.

7.2.3 Taylor Pricing Example

Here is an example where under Taylor contracts non-zero inflation eliminates price dispersion. All contracts last two periods so the price level and dispersion are given respectively by

$$P_t^{1-\theta} = \frac{1}{2}(p_t^*)^{1-\theta} + \frac{1}{2}(p_{t-1}^*)^{1-\theta}$$

$$\Delta_t = \frac{1}{2}\left(\frac{p_t^*}{P_t}\right)^{-\theta} + \frac{1}{2}\left(\frac{p_{t-1}^*}{P_t}\right)^{-\theta}$$

Suppose $\theta = 2$, $p_0^* = 1$, $p_1^* = 2$ which solves to give price level $P_1 = \frac{4}{3}$ and price dispersion $\Delta_1 = \frac{10}{9}$. Now consider time $t = 2$ the firms that set their price in period

0 now get to reset their price. Therefore the reset price $p_0^* = 1$ is replaced by p_2^* with $p_1^* = 2$ the other price in the economy. Now applying Lemma 2 $\Delta_2 = 1$ if and only if $p_2^* = 2$. This implies $P_2 = 2$ then inflation is non-zero in fact $\pi_t = 50\%$.

Bibliography

Daron Acemoglu. *Introduction to modern economic growth.* Princeton University Press, 2008.

Daron Acemoglu and Martin Kaae Jensen. Robust comparative statics in large static games. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 3133–3139. IEEE, 2010.

Daron Acemoglu, Martin Kaae Jensen, et al. Robust comparative statics in large dynamic economies. *Journal of Political Economy*, 123(3):587–640, 2015.

Daron Acemoglua and Martin Kaae Jensenb. Aggregate comparative statics. *Games and Economic Behavior*, 81:27–49, 2013.

Y Achdou, FJ Buera, JM Lasry, PL Lions, and B Moll. Partial differential equation models in macroeconomics. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 372(2028), 2014.

Klaus Adam. Optimal monetary policy with imperfect common knowledge. *Journal of Monetary Economics*, 54(2):267–301, 2007.

Klaus Adam and Roberto M Billi. Discretionary monetary policy and the zero lower bound on nominal interest rates. *Journal of monetary Economics*, 54(3):728–752, 2007.

Bernardino Adão, Isabel Correia, and Pedro Teles. Gaps and triangles. *The Review of Economic Studies*, 70(4):699–713, 2003.

S Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics*, 109(3):659–684, 1994.

James W Albrecht and Bo Axell. An equilibrium model of search unemployment. *Journal of Political Economy*, 92(5):824–40, 1984.

Syed Zahid Ali and Sajid Anwar. Can exchange rate pass-through explain the price puzzle? *Economics Letters*, 145:56–59, 2016.

Charalambos D Aliprantis and Kim Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Science & Business Media, 2006.

Charalambos D Aliprantis and Kim C Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer Science & Business Media, 2007.

Charalambos D Aliprantis and C Kim. Border. 2006. infinite dimensional analysis: A hitchhikers guide.

Fernando Alvarez, Francesco Lippi, et al. Price setting with menu cost for multiproduct firms. *Econometrica*, 82(1):89–135, 2014.

Fernando E Alvarez, Francesco Lippi, Luigi Paciello, et al. Phillips curves with observation and menu costs. *Journal of the European Economic Association forthcoming*, 2015.

Fernando E Alvarez, Francesco Lippi, Luigi Paciello, et al. Monetary shocks in models with inattentive producers. *Review of Economic Studies*, 83(2):421–459, 2016.

Luis Alvarez, Emmanuel Dhyne, Marco Hoeberichts, Claudia Kwapil, Hervé Le Bihan, Patrick Lünnemann, Fernando Martins, Roberto Sabbatini, Harald Stahl, Philip Vermeulen, et al. Sticky prices in the euro area: A summary of new micro-evidence. *Journal of the European Economic Association*, 4(2-3):575–584, 2006.

Sergio Afonso Lago Alves. Lack of divine coincidence in new keynesian models. *Journal of Monetary Economics*, 67:33–46, 2014.

Robert Amano and Malik Shukayev. Risk premium shocks and the zero bound on nominal interest rates. *Journal of Money, Credit and Banking*, 44(8):1475–1505, 2012.

Robert M Anderson et al. Non-standard analysis with applications to economics. *Handbook of mathematical economics*, 4:2145–2208, 1991.

Andrew Ang, Sen Dong, and Monika Piazzesi. No-arbitrage taylor rules. Technical report, National Bureau of Economic Research, 2007.

Aloísio Araújo, Susan Schommer, and Michael Woodford. Conventional and unconventional monetary policy with endogenous collateral constraints. Technical report, National Bureau of Economic Research, 2013.

Ludwig Arnold. *Random dynamical systems*. Springer Science & Business Media, 2013.

Kenneth J Arrow, Mordecai Kurz, et al. Optimal growth with irreversible investment in a ramsey model. *Econometrica*, 38(2):331–344, 1970.

Guido Ascari. Price/wage staggering and persistence: A unifying framework. *Journal of Economic Surveys*, 17(4):511–540, 2003.

Guido Ascari and Neil Rankin. Staggered wages and output dynamics under disinflation. *Journal of Economic Dynamics and Control*, 26(4):653–680, 2002.

Guido Ascari and Tiziano Ropele. Trend inflation, taylor principle, and indeterminacy. *Journal of Money, Credit & Banking*, 41(8):1557–1585, 2009.

Guido Ascari and Argia M Sbordone. The macroeconomics of trend inflation. *Journal of Economic Literature*, 52(3):679–739, 2014.

Guido Ascari, Louis Phaneuf, and Eric Sims. On the welfare and cyclical implications of moderate trend inflation. Technical report, National Bureau of Economic Research, Inc, 2015.

Michael J Ashton et al. "real-feel" inflation: Quantitative estimation of inflation perceptions. *Business Economics*, 47(1):14–26, 2012.

Susan Athey et al. Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187–223, 2002.

Adrien Auclert and Matthew Rognlie. Inequality and aggregate demand. Technical report, National Bureau of Economic Research, 2018.

Adrien Auclert et al. Monetary policy and the redistribution channel. Technical report, 2017.

Alan Auerbach and Yuriy Gorodnichenko. Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, 4(2):1–27, 2012.

David Autor, David Dorn, Lawrence F Katz, Christina Patterson, John Van Reenen, et al. The fall of the labor share and the rise of superstar firms. Technical report, National Bureau of Economic Research, Inc, 2017.

Laurence Ball. Short-run money demand. *Journal of Monetary Economics*, 59(7): 622–633, 2012.

Laurence Ball and David Romer. Real rigidities and the non-neutrality of money. *The Review of Economic Studies*, 57(2):183–203, 1990.

Laurence Ball, David Romer, et al. Sticky prices as coordination failure. *American Economic Review*, 81(3):539–52, 1991.

Simcha Barkai. Declining labor and capital shares. 2017.

William A Barnett. *Getting it wrong: how faulty monetary statistics undermine the fed, the financial system, and the economy.* MIT Press, 2012.

William A Barnett and Evgeniya A Duzhak. Empirical assessment of bifurcation regions within new keynesian models. *Economic Theory*, pages 99–128, 2010.

William A Barnett, Yijun He, et al. Stabilization policy as bifurcation selection: Would stabilization policy work if the economy really were unstable? *Macroeconomic Dynamics*, 6(05):713–747, 2002.

Robert Barro. Are government bonds net wealth? *Journal of Political Economy*, 1974.

Robert Barro. Rational expectations and the role of monetary policy. *Journal of Monetary Economics*, 1976.

Robert J Barro. A theory of monopolistic price adjustment. *The Review of Economic Studies*, pages 17–26, 1972.

Robert J Barro and Robert G King. Time-separable preferences and intertemporal-substitution models of business cycles. *The Quarterly Journal of Economics*, 99 (4):817–39, 1984.

Nicoletta Batini and Edward Nelson. The lag from monetary policy actions to inflation: Friedman revisited. *International Finance*, 4(3):381–400, 2001.

Marianne Baxter, Robert G King, et al. Measuring business cycles: Approximate band-pass filters for economic time series. *The Review of Economics and Statistics*, 81(4):575–593, 1999.

Paul Beaudry and Franck Portier. Stock prices, news, and economic fluctuations. *The American Economic Review*, pages 1293–1307, 2006.

Paul Beaudry and Franck Portier. When can changes in expectations cause business cycle fluctuations in neo-classical settings? *Journal of Economic Theory*, 135(1): 458–477, 2007.

Paul Beaudry, Dana Galizia, Franck Portier, et al. Reviving the limit cycle view of macroeconomic fluctuations. Technical report, National Bureau of Economic Research, Inc, 2015.

Gary S Becker et al. A theory of marriage: Part i. *Journal of Political Economy*, 81 (4):813–846, 1973.

Frederico Belo, Xiaoji Lin, and Santiago Bazdresch. Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy*, 122 (1):129–177, 2014.

Michael T Belongia and Peter N Ireland. Interest rates and money in the measurement of monetary policy. *Journal of Business & Economic Statistics*, (just-accepted): 00–00, 2014.

Jess Benhabib, Stephanie Schmitt-Grohé, and Martín Uribe. Chaotic interest-rate rules. *American Economic Review*, 92(2):72–78, 2002.

Pierpaolo Benigno, Gauti B Eggertsson, and Federica Romei. Dynamic debt deleveraging and optimal monetary policy. Technical report, National Bureau of Economic Research, 2014.

Deputy Governor Jarle Bergo. Interest rate projections in theory and practice1. *Economic Bulletin*, 1(0):7, 2007.

Ben S Bernanke and Mark Gertler. Should central banks respond to movements in asset prices? *American Economic Review*, 91(2):253–257, 2001.

Ben S Bernanke and Ilian Mihov. What does the bundesbank target? *European economic review*, 41(6):1025–1053, 1997.

Ben S Bernanke, Jean Boivin, Piotr Eliasz, et al. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *The Quarterly Journal of Economics*, 120(1):387–422, 2005.

Giuseppe Bertola, Aurelijus Dabusinskas, Marco Hoeberichts, Mario Izquierdo, Claudia Kwapil, Jeremi Montornès, and Daniel Radowski. Price, wage and employment response to shocks: evidence from the wdn survey. *Labour Economics*, 19(5):783–791, 2012.

Stephen Beveridge and Charles Nelson. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of thebusiness cycle'. *Journal of Monetary Economics*, 7(2):151–174, 1981.

Truman Bewley. The permanent income hypothesis: A theoretical formulation. *Journal of Economic Theory*, 16(2):252–292, 1977.

Roberto M Billi et al. Optimal inflation for the us economy. *American Economic Journal: Macroeconomics*, 3(3):29–52, 2011.

Mark Bils, Peter J Klenow, et al. Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985, 2004.

Olivier Blanchard and Jordi Galí. Real wage rigidities and the new keynesian model. *Journal of Money, Credit and Banking*, 39(s1):35–65, 2007.

Olivier Blanchard and Charles Kahn. The solution of linear difference models under rational expectations. *Econometrica*, 48(5):1305–11, 1980.

Olivier Blanchard and Nobuhiro Kiyotaki. Monopolistic competition and the effects of aggregate demand. *American Economic Review*, 77(4):647–66, 1987.

A Blanchet and G Carlier. From nash to cournot-nash equilibria via the monge-kantorovich problem. *Philosophical Transactions of the Royal Society of London Series A*, 372:20130398–20130398, 2014.

Adrien Blanchet and Guillaume Carlier. Optimal transport and cournot-nash equilibria. *Mathematics of Operations Research*, 41(1):125–145, 2016.

Alan S Blinder and Stanley Fischer. Inventories, rational expectations, and the business cycle. *Journal of Monetary Economics*, 8(3):277–304, 1981.

Alan S Blinder et al. Learning by asking those who are doing. *Eastern Economic Journal*, 16(4):297–306, 1990.

Alan S Blinder et al. Why are prices sticky? preliminary results from an interview study. *American Economic Review*, 81(2):89–96, 1991.

Alan S Blinder et al. Central-bank credibility: Why do we care? how do we build it? *American Economic Review*, 90(5):1421–1431, 2000.

Timo Boppart, Per Krusell, and Kurt Mitman. Exploiting mit shocks in heterogeneous-agent economies: The impulse response as a numerical derivative. Technical report, National Bureau of Economic Research, Inc, 2017.

Hafedh Bouakez, Emanuela Cardia, and Francisco J Ruge-Murcia. The transmission of monetary policy in a multisector economy. *International Economic Review*, 50 (4):1243–1266, 2009.

Hafedh Bouakez, Emanuela Cardia, and Francisco Ruge-Murcia. Sectoral price rigidity and aggregate dynamics. *European Economic Review*, 65:1–22, 2014.

Stephen P Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

R Braun and Tomoyuki Nakajima. Making the case for a low intertemporal elasticity of substitution. Technical report, Kyoto University, Institute of Economic Research, 2011.

Paulo B Brito, Luís F Costa, and Huw D Dixon. From sunspots to black holes: Singular dynamics in macroeconomic models. *Sunspots and Non-Linear Dynamics*, page 41, 2017.

William A Brock, Steven N Durlauf, and Kenneth D West. Model uncertainty and policy evaluation: Some theory and empirics'. *Journal of Econometrics*, 136(2): 629–664, 2007.

Tobias Broer, Niels-Jakob Harbo Hansen, Per Krusell, Erik Öberg, et al. The new keynesian transmission mechanism: A heterogenous-agent perspective. Technical report, CEPR Discussion Papers, 2016.

Francisco J Buera, Yongseok Shin, et al. Financial frictions and the persistence of history: A quantitative exploration. *Journal of Political Economy*, 121(2):221–272, 2013.

Kenneth Burdett and Dale T Mortensen. Wage differentials, employer size, and unemployment. *International Economic Review*, 39(2):257–73, 1998.

Kenneth Burdett, Kenneth L Judd, et al. Equilibrium price dispersion. *Econometrica*, 51(4):955–69, 1983.

Stephen Burgess, Emilio Fernandez-Corugedo, Charlotta Groth, Richard Harrison, Francesca Monti, Konstantinos Theodoridis, and Matt Waldron. The bank of england's forecasting platform: Compass, maps, ease and the suite of models. Technical report, Bank of England, 2013.

Arthur F Burns and Wesley C Mitchell. Measuring business cycles. Technical report, National Bureau of Economic Research, Inc, 1946.

Craig Burnside. Detrending and business cycle facts: A comment. *Journal of Monetary Economics*, 41:513–532, 1998.

A Colin Cameron, Severin Borenstein, and Richard J Gilbert. Do gasoline prices respond asymmetrically to crude oil price changes? *Quarterly Journal of Economics*, 112(1):305–39, 1997.

Fabio Canova. Detrending and turning points. *European Economic Review*, 38(3): 614–623, 1994.

Fabio Canova and Luca Sala. Back to square one: Identification issues in dsge models. *Journal of Monetary Economics*, 56(4):431–449, 2009.

Fabio Canova et al. Detrending and business cycle facts. *Journal of Monetary Economics*, 41(3):475–512, 1998.

H Carley and MD Taylor. A new proof of sklar's theorem. In *Distributions with given marginals and statistical modelling*, pages 29–34. Springer, 2002.

Guillaume Carlier, Victor Chernozhukov, Alfred Galichon, et al. Vector quantile regression: an optimal transport approach. *The Annals of Statistics*, 44(3):1165–1192, 2016.

Julio A Carrillo, Patrick Fève, and Julien Matheron. Monetary policy inertia or persistent shocks: A dsge analysis. *International Journal of Central Banking*, 3(2):1–38, 2007.

Christopher D Carroll, Miles S Kimball, et al. On the concavity of the consumption function. *Econometrica*, 64(4):981–992, 1996.

Carlos Carvalho. Heterogeneity in price stickiness and the real effects of monetary shocks. *Frontiers in Macroeconomics*, 6(3), 2006.

Efrem Castelnuovo and Paolo Surico. Monetary policy, inflation expectations and the price puzzle. *The Economic Journal*, 549(120):1262–1283, 2010.

Jagjit S Chadha, Luisa Corrado, and Sean Holly. A note on money and the conduct of monetary policy. *Macroeconomic Dynamics*, 18(08):1854–1883, 2014.

VV Chari, Patrick J Kehoe, and Ellen R McGrattan. New keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, 1(1):242–66, 2009.

Victor Chernozhukov, Han Hong, et al. Likelihood estimation and inference in a class of nonregular econometric models. *Econometrica*, 72(5):1445–1480, 2004.

Victor Chernozhukov, Alfred Galichon, Marc Hallin, Marc Henry, et al. Monge–kantorovich depth, quantiles, ranks and signs. *The Annals of Statistics*, 45(1):223–256, 2017.

Thanaset Chevapatrakul, Tae-Hwan Kim, Paul Mizen, et al. The taylor principle and monetary policy approaching a zero bound on nominal rates: Quantile regression results for the united states and japan. *Journal of Money, Credit and Banking*, 41 (8):1705–1723, 2009.

Pierre-Andr Chiappori, Robert McCann, Brendan Pass, et al. Multidimensional matching. Technical report, arXiv. org, 2016.

Pierre-André Chiappori, Robert J McCann, and Lars P Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Economic Theory*, 42(2):317–354, 2010.

Khai Xiang Chiong, Alfred Galichon, Matt Shum, et al. Duality in dynamic discrete-choice models. *Quantitative Economics*, 7(1):83–115, 2016.

Chi-Young Choi and Róisín O'Sullivan. Heterogeneous response of disaggregate inflation to monetary policy regime change: The role of price stickiness. *Journal of Economic Dynamics and Control*, 37(9):1814–1832, 2013.

M Ali Choudhary and Paul Levine. Risk-averse firms and employment dynamics. *Oxford Economic Papers*, 62(3):578–602, 2010.

Bent Jesper Christensen and Nicholas Kiefer. The exact likelihood function for an empirical job search model. *Econometric Theory*, 7(04):464–486, 1991.

Lawrence Christiano, Martin Eichenbaum, and Charles Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45, 2005.

Lawrence Christiano, Martin Eichenbaum, and Sergio Rebelo. When is the government spending multiplier large? *Journal of Political Economy*, 119(1):78–121, 2011.

Lawrence J Christiano, Martin Eichenbaum, Charles L Evans, et al. Monetary policy shocks: What have we learned and to what end? *Handbook of Macroeconomics*, 1: 65–148, 1999.

LawrenceJ Christiano and TerryJ Fitzgerald. The band pass filter. *International Economic Review*, 44(2):435–465, 2003.

Stijn Claessens, M Ayhan Kose, Marco E Terrones, et al. Financial cycles: What? how? when? In *NBER International Seminar on Macroeconomics*, volume 7, pages 303–344. University of Chicago Press, 2011.

Stijn Claessens, M Ayhan Kose, and Marco E Terrones. How do business and financial cycles interact? *Journal of International economics*, 87(1):178–190, 2012.

Richard Clarida, Jordi Galı, and Mark Gertler. Monetary policy rules in practice: some international evidence. *european economic review*, 42(6):1033–1067, 1998.

Edson A Coayla-Teran, Salah-Eldin A Mohammed, and Paulo Régis C Ruffino. Hartman-grobman theorems along hyperbolic stationary trajectories. *Discrete and Continuous Dynamical Systems*, 17(2):281, 2007.

Günter Coenen, Andrew T Levin, and Kai Christoffel. Identifying the influences of nominal and real rigidities in aggregate price-setting behavior. *Journal of Monetary Economics*, 54(8):2439–2466, 2007.

Nicolas Coeurdacier, Hélène Rey, and Pablo Winant. The risky steady state. *American Economic Review*, 101(3):398–401, 2011.

John F Cogan, Tobias Cwik, John B Taylor, and Volker Wieland. New keynesian versus old keynesian government spending multipliers. *Journal of Economic Dynamics and Control*, 34(3):281–295, 2010.

Timothy Cogley and James M Nason. Effects of the hodrick-prescott filter on trend and difference stationary time series implications for business cycle research. *Journal of Economic Dynamics and Control*, 1(19):253–278, 1995a.

Timothy Cogley and James M Nason. Output dynamics in real-business-cycle models. *The American Economic Review*, pages 492–511, 1995b.

Timothy Cogley and Argia M Sbordone. Trend inflation, indexation, and inflation persistence in the new keynesian phillips curve. *The American Economic Review*, 98(5):2101–2126, 2008.

Timothy Cogley, Giorgio E Primiceri, and Thomas J Sargent. Inflation-gap persistence in the us. *American Economic Journal: Macroeconomics*, 2(1):43–69, 2010.

Olivier Coibion. Testing the sticky information phillips curve. *The Review of Economics and Statistics*, 92(1):87–101, 2010.

Olivier Coibion. Are the effects of monetary policy shocks big or small? *American Economic Journal: Macroeconomics*, 4(2):1–32, 2012.

Olivier Coibion and Yuriy Gorodnichenko. Monetary policy, trend inflation, and the great moderation: An alternative interpretation. *American Economic Review*, 101 (1):341–70, 2011.

Olivier Coibion and Yuriy Gorodnichenko. Why are target interest rate changes so persistent? *American Economic Journal: Macroeconomics*, 4(4):126–162, 2012.

Olivier Coibion, Yuriy Gorodnichenko, and Johannes Wieland. The optimal inflation rate in new keynesian models: Should central banks raise their inflation targets in light of the zero lower bound? *The Review of Economic Studies*, 79(4):1371–1406, 2012.

Fabrice Collard and Harris Dellas. Tax distortions and the case for price stability. *Journal of Monetary Economics*, 52(1):249–273, 2005.

Luís Aguiar Conraria, Manuel MF Martins, and Maria Joana Soares. Estimating the taylor rule in the time-frequency domain. *NIPE Working Paper Series*, 18, 2014.

John B Conway. *A course in functional analysis*, volume 96. Springer Science & Business Media, 2013.

David Cook and Michael B Devereux. Cooperative fiscal policy at the zero lower bound. *Journal of the Japanese and international Economies*, 25(4):465–486, 2011a.

David Cook and Michael B Devereux. Optimal fiscal policy in a world liquidity trap. *European Economic Review*, 55(4):443–462, 2011b.

David Cook and Michael B Devereux. Sharing the burden: Monetary and fiscal responses to a world liquidity trap. *American Economic Journal: Macroeconomics*, 5(3):190–228, 2013.

Russell Cooper, Andrew John, et al. Coordinating coordination failures in keynesian models. *The Quarterly Journal of Economics*, 103(3):441–463, 1988.

Isabel Correia and Emmanuel Farhi. Unconventional fiscal policy at the zero bound. *American Economic Review*, 103(4):1172–1211, 2013.

Isabel Correia, Juan Pablo Nicolini, and Pedro Teles. Optimal fiscal and monetary policy: Equivalence results. *Journal of Political Economy*, 116(1):141–170, 2008.

Giancarlo Corsetti, Keith Kuester, André Meier, and Gernot J Müller. Debt consolidation and fiscal stabilization of deep recessions. *American Economic Review*, 100 (2):41–45, 2010.

Giancarlo Corsetti, Keith Kuester, André Meier, and Gernot J Müller. Sovereign risk, fiscal policy, and macroeconomic stability. *The Economic Journal*, 123(566): F99–F132, 2013.

John David Crawford. Introduction to bifurcation theory. *Reviews of Modern Physics*, 63(4):991, 1991.

Dean Croushore and Charles L Evans. Data revisions and the identification of monetary policy shocks. *Journal of Monetary Economics*, 53(6):1135–1160, 2006.

Francesco D'Acunto, Ryan Liu, Carolin Pflueger, Michael Weber, et al. Flexible prices and leverage. Technical report, CESifo Group Munich, 2017.

Tatiana Damjanovic and Charles Nolan. Relative price distortions and inflation persistence*. *The Economic Journal*, 120(547):1080–1099, 2010.

Carlos Viana de Carvalho. Aggregation and the ppp puzzle in a sticky-price model. *American Economic Review*, 101(6):2391–2424, 2011.

Oliver de Groot. Computing the risky steady state of dsge models. *Economics Letters*, 120(3):566–569, 2013.

Oliver de Groot. The risk channel of monetary policy. *International Journal of Central Banking*, 10(2):115–160, 2014.

Robert M de Jong, Neslihan Sakarya, et al. The econometrics of the hodrick-prescott filter. *The Review of Economics and Statistics*, 98(2):310–317, 2016.

Jan De Loecker and Jan Eeckhout. The rise of market power and the macroeconomic implications. Technical report, National Bureau of Economic Research, Inc, 2017.

Davide Debortoli and Jordi Galí. Monetary policy with heterogeneous agents: Insights from tank models. 2017.

Davide Debortoli, Jordi Galí, and Luca Gambetti. On the empirical (ir) relevance of the zero lower bound constraint. 2018.

Colin Decker, Elliott H Lieb, Robert J McCann, and Benjamin K Stephens. Unique equilibria and substitution effects in a stochastic model of the marriage market. *Journal of Economic Theory*, 148(2):778–792, 2013.

Fabio Del Missier, Rob Ranyard, and Nicolao Bonini. Perceived inflation: The role of product accessibility and attitudes towards inflation. *Journal of Economic Psychology*, 56(C):97–106, 2016.

Selva Demiralp, Kevin D Hoover, and Stephen J Perez. Still puzzling: evaluating the price puzzle in an empirically identified structural vector autoregression. *Empirical Economics*, 46(2):701–731, 2014.

Matthew Denes, Gauti B Eggertsson, and Sophia Gilbukh. Deficits, public debt dynamics and tax and spending multipliers. *The Economic Journal*, 123(566): F133–F163, 2013.

Richard Dennis. Consumption habits in a new keynesian business cycle model. *Journal of Money, Credit and Banking*, 41(5):1015–1030, 2009.

Alan K Detmeister, David E Lebow, Ekaterina V Peneva, et al. Inflation perceptions and inflation expectations. Technical report, Board of Governors of the Federal Reserve System (US), 2016.

Emmanuel Dhyne, Luis J Alvarez, Herve Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lunnemann, Fabio Rumler, and Jouko Vilmunen. Price changes in the euro area and the united states: Some facts from individual consumer price data. *Journal of Economic Perspectives*, 20(2):171–192, 2006.

William T Dickens, Lorenz Goette, Erica L Groshen, Steinar Holden, Julian Messina, Mark E Schweitzer, Jarkko Turunen, and Melanie E Ward. How wages change: Micro evidence from the international wage flexibility project. *Journal of Economic Perspectives*, 21(2):195–214, 2007.

Francis X Diebold, Georg Strasser, et al. On the correlation structure of microstructure noise: A financial economic approach. *Review of Economic Studies*, 80(4): 1304–1337, 2013.

Huw Dixon. A unified framework for using micro-data to compare dynamic time-dependent price-setting models. *The BE Journal of Macroeconomics*, 12(1), 2012.

Huw Dixon and Engin Kara. Can we explain inflation persistence in a way that is consistent with the microevidence on nominal rigidity? *Journal of Money, Credit and Banking*, 42(1):151–170, 2010.

Huw Dixon and Engin Kara. Contract length heterogeneity and the persistence of monetary shocks in a dynamic generalized taylor economy. *European Economic Review*, 55(2):280–292, 2011.

Huw Dixon and Hervé Le Bihan. Generalised taylor and generalised calvo price and wage setting: Micro-evidence with macro implications*. *The Economic Journal*, 122(560):532–554, 2012.

Stephen Donald and Harry Paarsch. Superconsistent estimation and inference in structural econometric models using extreme order statistics. *Journal of Econometrics*, 109(2):305–340, 2002.

Michael Dotsey and Robert G King. Pricing, production, and persistence. *Journal of the European Economic Association*, 4(5):893–928, 2006.

Randal Douc, Gersende Fort, Eric Moulines, Philippe Soulier, et al. Practical drift conditions for subgeometric rates of convergence. *The Annals of Applied Probability*, 14(3):1353–1377, 2004.

JM Dufour and C Hsiao. Identification. *The New Palgrave Dictionary of Economics, 2nd Edition. Palgrave Macmillan*, 2008.

Nelson Dunford and Jacob T Schwartz. Linear operators, vol. i. *Interscience, New York*, 1963, 1958.

Arnaud Dupuy and Alfred Galichon. Personality traits and the marriage market. *Journal of Political Economy*, 122(6):1271–1319, 2014.

Fabrizio Durante, Juan Fernandez-Sanchez, and Carlo Sempi. A topological proof of sklar's theorem. *Applied Mathematics Letters*, 26(9):945–948, 2013.

Stefania D'Amico, William English, David López-Salido, and Edward Nelson. The federal reserve's large-scale asset purchase programmes: Rationale and effects*. *The Economic Journal*, 122(564):F415–F446, 2012.

Rochelle Edge. The equivalence of wage and price staggering in monetary business cycle models. *Review of Economic Dynamics*, 5(3):559–585, 2002.

Gauti Eggertsson. The paradox of toil. Technical report, Federal Reserve Bank of New York, 2010.

Gauti Eggertsson, Andrea Ferrero, and Marco Del Negro. The response of monetary policy to financial distress. In *2009 Meeting Papers*, number 68. Society for Economic Dynamics, 2009.

Gauti Eggertsson, Andrea Ferrero, and Andrea Raffo. Can structural reforms help europe? *Journal of Monetary Economics*, 61:2–22, 2014.

Gauti Eggertsson, Federica Romei, Pierpaolo Benigno, et al. Dynamic debt deleveraging and optimal monetary policy. In *2015 Meeting Papers*, number 390. Society for Economic Dynamics, 2015.

Gauti B Eggertsson. Fiscal multipliers and policy coordination. *SSRN Working Paper Series*, 2006.

Gauti B Eggertsson. Great expectations and the end of the depression. *The American Economic Review*, 98(4):1476–1516, 2008.

Gauti B Eggertsson. What fiscal policy is effective at zero interest rates? In *NBER Macroeconomics Annual 2010, Volume 25*, pages 59–112. University of Chicago Press, 2011.

Gauti B Eggertsson. Was the new deal contractionary? *American Economic Review*, 102(1):524–55, 2012.

Gauti B Eggertsson and Marc P Giannoni. The inflation-output trade-off revisited. *FRB of New York Staff Report*, (608), 2013.

Gauti B Eggertsson and Paul R Krugman. Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *The Quarterly Journal of Economics*, 127(3): 1469–1513, 2012.

Gauti B Eggertsson and Neil R Mehrotra. A model of secular stagnation. Technical report, National Bureau of Economic Research, 2014.

Gauti B Eggertsson and Michael Woodford. Policy options in a liquidity trap. *American Economic Review*, pages 76–79, 2004.

Gauti B Eggertsson et al. Zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity*, 2003(1):139–233, 2003.

Martin Eichenbaum, Nir Jaimovich, and Sergio Rebelo. Reference prices, costs, and nominal rigidities. *American Economic Review*, 101(1):234–62, 2011.

Ivar Ekeland, Alfred Galichon, and Marc Henry. Comonotonic measures of multivariate risks. *Mathematical finance*, 22(1):109–132, 2012.

Martin Ellison and Thomas J Sargent. A defense of the fomc. *International Economic Review*, 53(4):1047–1065, 2012.

Martin Ellison and Andreas Tischbirek. Unconventional government debt purchases as a supplement to conventional monetary policy. *Journal of Economic Dynamics and Control*, 43:199–217, 2014.

Christopher Erceg and Jesper Lindé. Is there a fiscal free lunch in a liquidity trap? *Journal of the European Economic Association*, 12(1):73–107, 2014.

Neil R Ericsson, David F Hendry, and Kevin M Prestwich. The demand for broad money in the united kingdom, 1878–1993. *The Scandinavian Journal of Economics*, 100(1):289–324, 1998.

NEIL R Ericsson, DAVID F Hendry, and STEDMAN B Hood. Milton friedman as an empirical modeler. *Chapter*, 6:91–142, 2016.

Federico Etro and Lorenza Rossi. New-keynesian phillips curve with bertrand competition and endogenous entry. *Journal of Economic Dynamics and Control*, 51: 318–340, 2015.

George W Evans and Seppo Honkapohja. *Learning and expectations in macroeconomics*. Princeton University Press, 2001.

Emmanuel Farhi and Ivan Werning. Fiscal multipliers: Liquidity traps and currency unions. Technical report, National Bureau of Economic Research, 2012.

Emmanuel Farhi and Iván Werning. A theory of macroprudential policies in the presence of nominal rigidities. Technical report, National Bureau of Economic Research, 2013.

Emmanuel Farhi and Iván Werning. Labor mobility within currency unions. Technical report, National Bureau of Economic Research, 2014.

Mehdi Fatehi Nia and Fatemeh Rezaei. Hartman-grobman theorem for ifs. *arXiv preprint arXiv:1701.08495*, 2017.

Igor Fedotenkov. A bootstrap method to test for the existence of finite moments. *Journal of Nonparametric Statistics*, 25(2):315–322, 2013.

Igor Fedotenkov. A note on the bootstrap method for testing the existence of finite moments. *Statistica*, 74(4):447–453, 2014.

Jesús Fernández-Villaverde, Luis Garicano, and Tano Santos. Political credit cycles: The case of the eurozone. *Journal of Economic Perspectives*, 27(3):145–66, 2013.

Eric Fesselmeyer, Haoming Liu, and Alberto Salvo. How do households discount over centuries? evidence from singapore's private housing market. Technical report, Institute for the Study of Labor (IZA), 2016.

Alessio Figalli, Young-Heon Kim, and Robert J McCann. When is multidimensional screening a convex program? *Journal of Economic Theory*, 146(2):454–478, 2011.

Stanley Fischer. Long-term contracts, rational expectations, and the optimal money supply rule. *Journal of Political Economy*, 85(1):191–205, 1977.

Christopher J Flinn, James J Heckman, Thank G Chamberlain, T Coleman, B Jovanovic, B Miller, and James J Heclanan. New methods for analyzing structural models of labor force dynamics. In *Journal of Econometrics*. Citeseer, 1982.

Milton Friedman. The role of monetary policy. *The American Economic Review*, 58 (1):1–17, 1968a.

Milton Friedman. The role of monetary policy: Presidential address american economic association, 1967. *American Economic Review*, 58(1):1–17, 1968b.

Milton Friedman and Anna Schwartz. Alternative approaches to analyzing economic data. *American Economic Review*, 81(1):39–49, 1991.

Roman Frigg, Joseph Berkovitz, and Fred Kronz. The ergodic hierarchy. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2016 edition, 2016.

Drew Fudenberg and Kevin He. Learning and equilibrium refinements in signalling games. Technical report, arXiv. org, 2017a.

Drew Fudenberg and Kevin He. Player-compatible equilibrium. Technical report, arXiv. org, 2017b.

Jeff Fuhrer and Geoff Tootell. Eyes on the prize: How did the fed respond to the stock market? *Journal of Monetary Economics*, 55(4):796–805, 2008.

Jeffrey C Fuhrer. Habit formation in consumption and its implications for monetary-policy models. *American Economic Review*, pages 367–390, 2000.

Jeffrey C Fuhrer. Intrinsic and inherited inflation persistence. *International Journal of Central Banking*, 2006.

Xavier Gabaix, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. The dynamics of inequality. *Econometrica*, 6(84):2071–2111, 2016.

Jordi Galı and Mark Gertler. Inflation dynamics: A structural econometric analysis. *Journal of monetary Economics*, 44(2):195–222, 1999.

Alfred Galichon. *Optimal Transport Methods in Economics.* Princeton University Press, 2016.

Alfred Galichon. A survey of some recent applications of optimal transport methods to econometrics. *The Econometrics Journal*, 20(2), 2017.

Alfred Galichon and Bernard Salanie. Cupid's invisible hand: Social surplus and identification in matching models. 2015.

Stefan Gerlach and Lars EO Svensson. Money and inflation in the euro area: A case for monetary indicators? *Journal of Monetary Economics*, 50(8):1649–1672, 2003.

Stefan Gerlach and Peter Tillmann. Inflation targeting and inflation persistence in asia–pacific. *Journal of Asian Economics*, 23(4):360–373, 2012.

Mark Gertler and Peter Karadi. A model of unconventional monetary policy. *Journal of Monetary Economics*, 1(58):17–34, 2011.

Mark Gertler and Peter Karadi. Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76, 2015.

Mark Gertler and John Leahy. A phillips curve with an ss foundation. *Journal of Political Economy*, 116(3), 2008.

Mark Gertlera and Peter Karadib. Qe 1 vs. 2 vs. 3...: A framework for analyzing large-scale asset purchases as a monetary policy tool. *International Journal of Central Banking*, 2013.

Sebastian Giesen, Oliver Holtemöller, Juliane Scharff, and Rolf Scheufele. The halle economic projection model. *Economic Modelling*, 29(4):1461–1472, 2012.

Stefano Giglio, Matteo Maggiori, and Johannes Stroebel. Editor's Choice Very Long-Run Discount Rates. *The Quarterly Journal of Economics*, 130(1):1–53, 2015.

Stefano Giglio, Matteo Maggiori, Johannes Stroebel, et al. No-bubble condition: Model-free tests in housing markets. *Econometrica*, 84:1047–1091, 2016.

Simon Gilchrist, David López-Salido, and Egon Zakrajšek. Monetary policy and real borrowing costs at the zero lower bound. *American Economic Journal: Macroeconomics*, 7(1):77–109, 2015.

Gregory Givens. Deep or aggregate habit formation? evidence from a new-keynesian business cycle model. Technical report, University Library of Munich, Germany, 2013.

Marvin Goodfriend and Robert King. The new neoclassical synthesis and the role of monetary policy. In *NBER Macroeconomics Annual 1997, Volume 12*, pages 231–296. MIT Press, 1997.

Charles Goodhart. The evolution of central banks. 1988.

Yuriy Gorodnichenko. Endogenous information, menu costs and inflation persistence. Technical report, National Bureau of Economic Research, 2008.

Yuriy Gorodnichenko and Michael Weber. Are sticky prices costly? evidence from the stock market. *American Economic Review*, 106(1):165–99, 2016.

Jean-Michel Grandmont. Temporary equilibrium. Technical report, Centre de Recherche en Economie et Statistique, 2006.

WH Greene. Econometric analysis. 2003.

Michael Greinecker and Christopher Kah. Pairwise stable matching in large economies. Technical report, University of Graz, Department of Economics, 2018.

Bernardo Guimaraes and Kevin D Sheedy. Sales and monetary policy. *The American Economic Review*, 101(2):844–876, 2011.

German Gutierrez. Investigating global labor and profit shares. *Unpublished Working Paper, SSRN*, 2017.

FH Hahn et al. A stable adjustment process for a competitive economy. *Review of Economic Studies*, 29(1):62–65, 1961.

Frank Horace Hahn. Keynesian economics and general equilibrium theory: reflections on some current debates. In *The microeconomic foundations of macroeconomics*, pages 25–40. Springer, 1977.

Martin Hairer. Ergodic properties of markov processes. 2006.

Paul R Halmos. Finite-dimensional vector spaces. the university series in undergraduate mathematics, 1958.

Paul R Halmos. *Measure theory*, volume 18. Springer, 2013.

James D Hamilton. Why you should never use the hodrick-prescott filter. Technical report, National Bureau of Economic Research, Inc, 2017.

James D Hamilton, Seth Pruitt, and Scott Borger. Estimating the market-perceived monetary policy rule. *American Economic Journal. Macroeconomics*, 3(3):1, 2011.

James Douglas Hamilton. *Time series analysis*, volume 2.

Lars Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–54, 1982.

Lars Peter Hansen and Kenneth J Singleton. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica*, 50(5):1269–1286, 1982.

AC Harvey, A Jaeger, et al. Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics*, 8(3):231–47, 1993.

John Hassler, Petter Lundvik, and Torsten Persson. *The Swedish business cycle: stylized facts over 130 years.* Number 22. 1992.

Tomas Havranek and Marek Rusnak. Transmission Lags of Monetary Policy: A Meta-Analysis. *International Journal of Central Banking*, 9(4):39–76, December 2013.

Tomáš Havránek et al. Measuring intertemporal substitution: The importance of method choices and selective reporting. *Journal of the European Economic Association*, 13(6):1180–1204, 2015.

FA Hayek. The use of knowledge in society. *The American Economic Review*, 35(4): 519–530, 1945.

James J Heckman, Rosa L Matzkin, Lars Nesheim, et al. Nonparametric identification and estimation of nonadditive hedonic models. *Econometrica*, 78(5):1569–1591, 2010.

Dale W Henderson and Warwick J McKibbin. A comparison of some basic monetary policy regimes for open economies: implications of different degrees of instrument adjustment and wage persistence. In *Carnegie-Rochester Conference Series on Public Policy*, volume 39, pages 221–317. Elsevier, 1993.

David Hendry and Neil Ericsson. An econometric analysis of uk money demand in'monetary trends in the united states and the united kingdom'by milton friedman and anna schwartz. *American Economic Review*, 81(1):8–38, 1991.

JR Hicks. Mr. keynes and the" classics"; a suggested interpretation. *Econometrica*, 5 (2):147–159, 1937.

Keisuke Hirano and Jack R Porter. Asymptotic efficiency in parametric structural models with parameter-dependent support. *Econometrica*, 71(5):1307–1338, 2003.

Bart Hobijn, Federico Ravenna, and Andrea Tambalotti. Menu costs at work: Restaurant prices and the introduction of the euro. *The Quarterly Journal of Economics*, 121(3):1103–1131, 2006.

Robert J Hodrick and Edward C Prescott. Postwar us business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16, 1997.

A Holmsen, JF Qvigstad, and Øistein Røisland. Implementing and communicating optimal monetary policy. *Norges Bank Staff Memo (Monetary Policy) No*, 3, 2007.

Stefan Homburg. *A Study in Monetary Macroeconomics*. Oxford University Press, 2017.

Hugo Hopenhayn and Edward Prescott. Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica*, 60(6):1387–406, 1992.

Odilo W Huber. Frequency of price increases and perceived inflation. an experimental investigation. *Journal of Economic Psychology*, 32(5):651–661, 2011.

Mark Huggett et al. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*, 17(5-6):953–969, 1993.

Daisuke Ikeda. Optimal inflation rates with the trending relative price of investment. *Journal of Economic Dynamics and Control*, 56:20–33, 2015.

Jean Imbs, Eric Jondeau, and Florian Pelgrin. Sectoral phillips curves and the aggregate phillips curve. *Journal of Monetary Economics*, 58(4):328–344, 2011.

Yuko Imura, Julia Thomas, et al. International transmission of credit shocks in an equilibrium model with production heterogeneity. Technical report, Bank of Canada, 2015.

Peter N Ireland. On the welfare cost of inflation and the recent behavior of money demand. *American Economic Review*, 99(3):1040–1052, 2009.

Nir Jaimovich and Sergio Rebelo. Behavioral theories of the business cycle. *Journal of the European Economic Association*, 5(2-3):361–368, 2007.

Nir Jaimovich and Sergio Rebelo. Can news about the future drive the business cycle? *The American Economic Review*, 99(4):1097–1118, 2009.

Robert A Jarrow. The zero-lower bound on interest rates: Myth or reality? *Finance Research Letters*, 10(4):151–156, 2013.

Martin Kaae Jensen. Distributional comparative statics. *The Review of Economic Studies*, 85(1):581–610, 2017.

Gabriel Jiménez, Steven Ongena, José-Luis Peydró, and Jesús Saurina. Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking? *Econometrica*, 82(2):463–505, 2014.

Lei Jin, Suojin Wang, and Haiyan Wang. A new non-parametric stationarity test of time series in the time domain. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(5):893–922, 2015.

Lars Jonung et al. Perceived and expected rates of inflation in sweden. *American Economic Review*, 71(5):961–68, 1981.

Kenneth L Judd. *Numerical methods in economics.* MIT press, 1998.

Michel Juillard and Ondra Kamenik. Solving sdge models: approximation about the stochastic steady state. Technical report, Society for Computational Economics, 2005.

Gunes Kamber, Chris McDonald, Nick Sander, and Konstantinos Theodoridis. Modelling the business cycle of a small open economy: The reserve bank of new zealand's dsge model. *Economic Modelling*, 59:546–569, 2016.

George Kapetanios, Haroon Mumtaz, Ibrahim Stevens, and Konstantinos Theodoridis. Assessing the economy-wide effects of quantitative easing*. *The Economic Journal*, 122(564):F316–F347, 2012.

Greg Kaplan, Benjamin Moll, and Giovanni L Violante. Monetary policy according to hank. *American Economic Review*, 108(3):697–743, 2018.

Daniel Kaufmann and Sarah Lein. Is there a swiss price puzzle? *Swiss Journal of Economics and Statistics (SJES)*, 148(I):57–75, 2012.

Phillip Kearns, Adrian Pagan, et al. Estimating the density tail index for financial time series. *The Review of Economics and Statistics*, 79(2):171–175, 1997.

Patrick J Kehoe and Virgiliu Midrigin. Temporary price changes and the real effects of monetary policy. *NBER Working Paper*, (W14392), 2008.

John Maynard Keynes. The general theory of interest, employment and money, 1936.

Aubhik Khan, Robert G King, and Alexander L Wolman. Optimal monetary policy. *The Review of Economic Studies*, 70(4):825–860, 2003.

Aubhik Khan, Julia K Thomas, et al. Credit shocks and aggregate fluctuations in an economy with production heterogeneity. *Journal of Political Economy*, 121(6): 1055–1107, 2013.

Aubhik Khan et al. Aggregate fluctuations in a quantitative overlapping generations economy with unemployment risk. In *2016 Meeting Papers*, number 1468. Society for Economic Dynamics, 2016.

Heejeong Kim and Aubhik Khan. Segmented Asset Markets and the Distribution of Wealth. Technical report, 2015.

Jinill Kim, Andrew Levin, and Tack Yun. Bifurcation in Perturbation Analysis:Calvo Pricing Examples. *Computational Economics*, 37(3):221–236, March

2011. doi: 10.1007/s10614-008-9155-1. URL https://ideas.repec.org/a/kap/compec/v37y2011i3p221-236.html.

Mervyn King. Changes in uk monetary policy: Rules and discretion in practice. *Journal of Monetary Economics*, 39(1):81–97, 1997.

Robert King and Sergio Rebelo. Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control*, 17(1-2):207–231, 1993.

Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica*, 46(1): 33–50, 1978.

Leonid Kogan. An equilibrium model of irreversible investment. *Journal of Financial Economics*, 62(2):201–245, 2001.

Donald Kohn. Policy challenges for the federal reserve. URL http://www.federalreserve.gov/newsevents/speech/kohn20091116a.htm. Kellogg Distinguished Lecture Series, Kellogg School of Management, Northwestern University, Evanston, Illinois.

Petar V Kokotovic, John J Allemong, James R Winkelman, and Joe H Chow. Singular perturbation and iterative separation of time scales. *Automatica*, 16(1):23–33, 1980.

Gary Koop, M Hashem Pesaran, and Ron P Smith. On identification of bayesian dsge models. *Journal of Business & Economic Statistics*, 31(3):300–314, 2013.

Anton Korinek and Alp Simsek. Liquidity trap and excessive leverage. Technical report, National Bureau of Economic Research, 2014.

Georgios P Kouretas and Mark E Wohar. The dynamics of inflation: a study of a large number of countries. *Applied Economics*, 44(16):2001–2026, 2012.

Andros Kourtellos and Thanasis Stengos. Varying coefficient models. In *Macroeconometrics and Time Series Analysis*, pages 372–377. Springer, 2010.

David M Kreps, Robert Wilson, et al. Sequential equilibria. *Econometrica*, 50(4): 863–894, 1982.

Dejan Krusec. The “price puzzle” in the monetary transmission vars with long-run restrictions. *Economics Letters*, 106(3):147–150, 2010.

Per Krusell and Anthony A Smith Jr. Income and wealth heterogeneity in the macroeconomy. *The Journal of Political Economy*, 106(5):867–896, 1998.

Michael Kumhof, Ricardo Nunes, Irina Yakadina, et al. Simple monetary rules under fiscal dominance. *Journal of Money, Credit and Banking*, 42(1):63–92, 2010.

Philip B Kurland, Ralph Lerner, et al. The founders' constitution. 1987.

Yuri Kuznetsov. *Elements of Applied Bifurcation Theory*, volume 112. Springer Science & Business Media, 2013.

Finn E Kydland and Edward C Prescott. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370, 1982.

Ricardo Lagos and Randall Wright. A unified framework for monetary theory and policy analysis. *Journal of Political Economy*, 113:463–484, 2005.

Oliver Landmann. Rotating slumps in a monetary union. *Open Economies Review*, 23(2):303–317, 2012.

Serge Lang. *Fundamentals of Differential Geometry*, volume 191. Springer Science & Business Media, 2001.

Jennifer La'O. A traffic jam theory of recessions. Technical report, Society for Economic Dynamics, 2013.

Andrzej Lasota and Michael C Mackey. *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*, volume 97. Springer Science & Business Media, 1998.

Kevin Lee, M Hashem Pesaran, Ron Smith, et al. Growth and convergence in multi-country empirical stochastic solow model. *Journal of Applied Econometrics*, 12(4): 357–92, 1997.

Erich Leo Lehmann and George Casella. *Theory of point estimation*, volume 31. Springer Science & Business Media, 1998.

Stanford E Lehmberg. Parliamentary attainder in the reign of henry viii. *The Historical Journal*, 18(4):675–702, 1975.

Axel Leijonhufvud. *On Keynesian economics and the economics of Keynes*, volume 158. Oxford University Press New York, 1968.

Axel Leijonhufvud. *Information and coordination: essays in macroeconomic theory.* Oxford Univ Pr, 1981.

Andrew Levin, Volker Wieland, and John C Williams. The performance of forecast-based monetary policy rules under model uncertainty. *American Economic Review*, 93(3):622–645, 2003.

Andrew Levin, Volker Wieland, and John C Williams. 6 robustness of simple monetary policy rules under model uncertainty. *Monetary Policy Rules*, 31:263, 2007.

Andrew T Levin, Volker Wieland, John Williams, et al. Robustness of simple monetary policy rules under model uncertainty. *NBER Chapters*, pages 263–318, 1999.

Andrew T Levin, Alexei Onatski, John Williams, Noah M Williams, et al. Monetary policy under uncertainty in micro-founded macroeconometric models. *NBER Chapters*, pages 229–312, 2006.

Daniel Levy et al. The magnitude of menu costs: Direct evidence from large us supermarket chains. *The Quarterly Journal of Economics*, 112(3):791–825, 1997.

Vivien Lewis and Céline Poilly. Firm entry, markups and the monetary transmission mechanism. *Journal of Monetary Economics*, 59(7):670–685, 2012.

Vivien Lewis and Arnoud Stevens. Entry and markup dynamics in an estimated business cycle model. *European Economic Review*, 74:14–35, 2015.

Assar Lindbeck. Macroeconomic theory and the labor market. *European Economic Review*, 36(2):209–235, 1992.

Shiqing Ling and Michael McAleer. Stationarity and the existence of moments of a family of garch processes. *Journal of Econometrics*, 1(106):109–117, 2002.

John B Long Jr and Charles I Plosser. Real business cycles. *Journal of Political Economy*, 91(1):39–69, 1983.

Guido Lorenzoni. A theory of demand shocks. *American Economic Review*, 99(5):2050–84, 2009.

Guido Lorenzoni. Optimal monetary policy with uncertain fundamentals and dispersed information. *The Review of Economic Studies*, 77(1):305–338, 2010.

Guido Lorenzoni and Veronica Guerrieri. Credit crises, precautionary savings and the liquidity trap. In *2011 Meeting Papers*, number 1414. Society for Economic Dynamics, 2011.

Qin Lu, Yaakov Bar-Shalom, Peter Willett, Francesco Palmieri, and Fred Daum. The multidimensional cramér–rao–leibniz lower bound for likelihood functions with parameter-dependent support. *IEEE Transactions on Aerospace and Electronic Systems*, 53(5):2331–2343, 2017.

Robert E Lucas. Expectations and the neutrality of money. *Journal of economic theory*, 4(2):103–124, 1972.

Robert E. Lucas and Juan Pablo Nicolini. On the stability of money demand. Working papers, Federal Reserve Bank of Minneapolis, February 2015.

Robert E Lucas Jr. Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. Elsevier, 1976.

Celsa Machado. Measuring business cycles: The real business cycle approach and related controversies. Technical report, Universidade do Porto, Faculdade de Economia do Porto, 2001.

Bartosz Maćkowiak and Mirko Wiederholt. Business cycle dynamics under rational inattention. *The Review of Economic Studies*, page rdv027, 2015.

Bartosz Mackowiak, Mirko Wiederholt, et al. Business cycle dynamics under rational inattention. In *2008 Meeting Papers*, number 1059. Society for Economic Dynamics, 2008.

Benjamin A Malin, Eric Anderson, Emi Nakamura, Duncan Simester, and Jon Steinsson. Informational rigidities and the stickiness of temporary sales. Technical report, Federal Reserve Bank of Minneapolis, 2015.

N Gregory Mankiw. Small menu costs and large business cycles: A macroeconomic model of monopoly. *The Quarterly Journal of Economics*, pages 529–537, 1985.

N Gregory Mankiw and Ricardo Reis. Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328, 2002.

N Gregory Mankiw and Ricardo Reis. Pervasive stickiness. *American Economic Review*, 96(2):164–169, 2006.

N Gregory Mankiw and Ricardo Reis. Sticky information in general equilibrium. *Journal of the European Economic Association*, 5(2-3):603–613, 2007.

N Gregory Mankiw and Matthew Weinzierl. The optimal taxation of height: A case study of utilitarian income redistribution. *American Economic Journal. Economic Policy*, 2(1):155, 2010.

Henry B Mann and Abraham Wald. On stochastic limit and order relationships. *The Annals of Mathematical Statistics*, 14(3):217–226, 1943.

Alfred Marshall. Principles of political economy. *Maxmillan, New York*, 1890.

Rosa L Matzkin et al. Nonparametric estimation of nonadditive random functions. *Econometrica*, 71(5):1339–1375, 2003.

Sophocles Mavroeidis. Weak identification of forward-looking models in monetary economics*. *Oxford Bulletin of Economics and statistics*, 66(s1):609–635, 2004.

Bennett McCallum. Robustness properties of a rule for monetary policy. In *Carnegie-Rochester Conference Series on Public Policy*, volume 29, pages 173–203. Elsevier, 1988.

Robert J McCann and Kelvin Shuangjian Zhang. On concavity of the monopolist's problem facing consumers with nonlinear price preferences. *arXiv preprint arXiv:1709.00632*, 2017.

Quasi-Arithmetic Means and Symmetric Polynomial Means. Handbook of means and their inequalities. *Mathematics and Its Applications*, 560, 2003.

Barbara Meller and Dieter Nautz. Inflation persistence in the euro area before and after the european monetary union. *Economic Modelling*, 29(4):1170–1176, 2012.

Laurence H Meyer, Eric T Swanson, Volker W Wieland, et al. Nairu uncertainty and nonlinear policy rules. *American Economic Review*, 91(2):226–231, 2001.

Sean P Meyn and Richard L Tweedie. Markov chains and stochastic stability. 2009.

Woodford Michael. Inflation stabilization and welfare. *The BE Journal of Macroeconomics*, 2(1):1–53, 2002.

Juillard Michel. Local approximation of dsge models around the risky steady state. Technical report, Department of Communication, University of Teramo, 2011.

Paul Milgrom and Joshua Mollner. Equilibrium selection in auctions and high stakes games. *Econometrica*, 86(1):219–261, 2018.

Benjamin Moll. Productivity losses from financial frictions: Can self-financing undo capital misallocation? *American Economic Review*, 104(10):3186–3221, 2014.

David S Moore, MC Spruill, et al. Unified large-sample theory of general chi-squared statistics for tests of fit. *The Annals of Statistics*, 3(3):599–616, 1975.

Roger B Myerson. Refinements of the nash equilibrium concept. *International Journal of Game Theory*, 7(2):73–80, 1978.

Seunghoon Na, Stephanie Schmitt-Grohé, Martin Uribe, and Vivian Z Yue. A model of the twin ds: Optimal default and devaluation. Technical report, National Bureau of Economic Research, Inc, 2014.

Emi Nakamura and Jon Steinsson. Five facts about prices: A reevaluation of menu cost models. *The Quarterly Journal of Economics*, 123(4):1415–1464, 2008.

Emi Nakamura and Jón Steinsson. Price setting in forward-looking customer markets. *Journal of Monetary Economics*, 58(3):220–233, 2011.

Emi Nakamura and Jon Steinsson. Price rigidity: Microeconomic evidence and macroeconomic implications. *Annual Review of Economics*, 5:133–163, 2013.

Roger B Nelsen. *An Introduction to Copulas*. Springer Science & Business Media, 2006.

Benjamin Nelson, Gabor Pinter, and Konstantinos Theodoridis. Do contractionary monetary policy shocks expand shadow banking? 2015.

Matthias Neuenkirch and Peter Tillmann. Inflation targeting, credibility, and non-linear Taylor rules. *Journal of International Money and Finance*, 41(C):30–45, 2014.

Whitney K Newey and Daniel McFadden. Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4:2111–2245, 1994.

Kristoffer Nimark. Dynamic pricing and imperfect common knowledge. *Journal of Monetary Economics*, 55(2):365–382, 2008.

Kristoffer Nimark. Man-bites-dog business cycles. *The American Economic Review*, 104(8):2320–2367, 2014.

Kazuo Nishimura and John Stachurski. Stability of stochastic optimal growth models: A new approach. *Journal of Economic Theory*, 122:289, 2005.

Georg Noldeke, Larry Samuelson, et al. The implementation duality. Technical report, Cowles Foundation for Research in Economics, Yale University, 2017.

Carlos Noton. Structural estimation of price adjustment costs in the european car market. *International Journal of Industrial Organization*, 49(November 2016):105–147, 2016.

Esa Nummelin and Pekka Tuominen. The rate of convergence in orey's theorem for harris recurrent markov chains with applications to renewal theory. *Stochastic Processes and Their Applications*, 15(3):295–311, 1983.

Giovanni Olivei and Silvana Tenreyro. The timing of monetary policy shocks. *American Economic Review*, 97(3):636–663, 2007.

Giovanni Olivei and Silvana Tenreyro. Wage-setting patterns and monetary policy: International evidence. *Journal of Monetary Economics*, 57(7):785–802, 2010.

A Onatski and J Stock. Robust monetary policy under model uncertainty in a small model of the us economy. *Macroeconomic Dynamics*, 6:85–110, 2002.

Gerard O'Reilly and Karl Whelan. Has euro-area inflation persistence changed over time? *Review of Economics and Statistics*, 87(4):709–720, 2005.

Lucjan T Orlowski. Ben s. bernanke, thomas laubach, frederic s. mishkin, and adam s. posen, inflation targeting: Lessons from the international experience., 2000.

Athanasios Orphanidesa and Volker Wielandb. Complexity and monetary policy. *International Journal of Central Banking*, 2013.

Ali Ozdagli and Michael Weber. Monetary policy through production networks: Evidence from the stock market. Technical report, National Bureau of Economic Research, Inc, 2017.

BG Pachpatte. Inequalities for differential and integral equations. 1997.

Luigi Paciello and Mirko Wiederholt. Exogenous information, endogenous information, and optimal monetary policy. *Review of Economic Studies*, 81(1), 2014.

Brendan Pass. Convexity and multi-dimensional screening for spaces with different dimensions. *Journal of Economic Theory*, 147(6):2399–2418, 2012.

DS Mitrinovic JE Pecaric and AM Fink. Classical and new inequalities in analysis, 1993.

Gerlach-Kristen Petra. Interest-rate smoothing: Monetary policy inertia or unobserved variables? *The BE Journal of Macroeconomics*, 4(1):1–19, 2004.

Edmund Phelps. Money-wage dynamics and labor-market equilibrium. *Journal of Political Economy*, 76, 1968.

Frederic Pivetta and Ricardo Reis. The persistence of inflation in the united states. *Journal of Economic dynamics and control*, 31(4):1326–1358, 2007.

Vasilios Plakandaras, Periklis Gogas, Rangan Gupta, and Theophilos Papadimitriou. Us inflation dynamics on long range data. Technical report, 2014.

Dario Pontiggia. Optimal long-run inflation and the new keynesian model. *Journal of Macroeconomics*, 34(4):1077–1094, 2012.

Adam S Posen. Why central banks should not burst bubbles*. *International Finance*, 9(1):109–124, 2006.

Federico Ravenna and Carl E Walsh. Optimal monetary policy with the cost channel. *Journal of Monetary Economics*, 53(2):199–216, 2006.

Doron Ravid. Bargaining with rational inattention. 2017.

Morten Ravn and Harald Uhlig. On adjusting the hodrick-prescott filter for the frequency of observations. *The Review of Economics and Statistics*, 84(2):371–375, 2002.

Morten O Ravn, Stephanie Schmitt-Grohė, Lenno Uuskula, et al. Deep habits and the dynamic effects of monetary policy shocks. *Journal of the Japanese and International Economies*, 24(2):236–258, 2010.

William H Rehnquist. *The Supreme Court*. Vintage, 2007.

Adam Reiff, Peter Karadi, et al. Menu costs, aggregate fluctuations and large shocks. In *2014 Meeting Papers*, number 914. Society for Economic Dynamics, 2014.

Carmen M Reinhart, Kenneth S Rogoff, et al. A decade of debt. *Central Banking, Analysis, and Economic Policies Book Series*, 18:97–135, 2014.

CM Reinhart. The return of financial repression. *Financial Stability Review*, (16): 37–48, 2012.

Ricardo Reis. Inattentive producers. *The Review of Economic Studies*, 73(256):793, 2006.

Elke Renner and Jean-Robert Tyran. Price rigidity in customer markets. *Journal of Economic Behavior & Organization*, 55(4):575–593, 2004.

John M Roberts. New keynesian economics and the phillips curve. *Journal of Money, Credit and Banking*, 27(4):975–84, 1995.

John M Roberts. Is inflation sticky? *Journal of Monetary Economics*, 39(2):173–196, 1997.

Matthew Rognlie, Andrei Shleifer, and Alp Simsek. Investment hangover and the great recession. Technical report, National Bureau of Economic Research, 2014.

Christina D Romer and David H Romer. A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084, 2004.

Christina D Romer and David H Romer. The fomc versus the staff: Where can monetary policymakers add value? *American Economic Review*, 98(2):230–35, 2008.

David Romer. Advanced macroeconomics, 2012.

Sheldon M Ross. Introduction to probability models. 2010.

Julio J Rotemberg. Monopolistic price adjustment and aggregate output. *The Review of Economic Studies*, 49(4):517–531, 1982.

Julio J Rotemberg. Customer anger at price increases, changes in the frequency of price adjustment and monetary policy. *Journal of Monetary Economics*, 52(4): 829–852, 2005.

Julio J Rotemberg. The federal reserve's abandonment of its 1923 principles. Technical report, National Bureau of Economic Research, 2014.

Julio J Rotemberg and Michael Woodford. Real-business-cycle models and the forecastable movements in output, hours, and consumption. *American Economic Review*, 86(1):71–89, 1996.

Julio J Rotemberg et al. Fair pricing. *Journal of the European Economic Association*, 9(5):952–981, 2011.

H Royden and P Fitzpatrick. Real analysis. 2010.

Julie Rozenberg, Adrien Vogt-Schilb, and Stephane Hallegatte. Transition to clean capital, irreversible investment and stranded assets. Policy Research Working Paper Series 6859, The World Bank, May 2014. URL https://ideas.repec.org/p/wbk/wbrwps/6859.html.

Jeremy Rudd and Karl Whelan. New tests of the new-keynesian phillips curve. *Journal of Monetary Economics*, 52(6):1167–1181, 2005.

Glenn Rudebusch. Is the fed too timid? monetary policy in an uncertain world. *The Review of Economics and Statistics*, 83(2):203–217, 2001.

Glenn D Rudebusch. Term structure evidence on interest rate smoothing and monetary policy inertia. *Journal of monetary economics*, 49(6):1161–1187, 2002.

Glenn D Rudebusch. Monetary policy inertia: Fact or fiction? *International Journal of Central Banking*, 2(4), 2006.

Francisco J Ruge-Murcia. Inflation targeting under asymmetric preferences. *Journal of Money, Credit & Banking (Ohio State University Press)*, 35(5), 2003.

Fabio Rumler. Estimates of the open economy new keynesian phillips curve for euro area countries. *open economies review*, 18(4):427–451, 2007.

Ludger Rüschendorf. On the distributional transform, sklar's theorem, and the empirical copula process. *Journal of Statistical Planning and Inference*, 139(11):3921–3927, 2009.

Marek Rusnák, Tomas Havranek, and Roman Horváth. How to solve the price puzzle? a meta-analysis. *Journal of Money, Credit and Banking*, 45(1):37–70, 2013.

Luca Sala, Ulf Söderström, Antonella Trigari, et al. Monetary policy under uncertainty in an estimated model with labor market frictions. *Journal of Monetary Economics*, 55(5):983–1006, 2008.

Paul A Samuelson. Economics-an introductory analysis, 1951.

Yuliy Sannikov. Contracts: The theory of dynamic principal-agent relationships and the continuous-time approach. In *10th World Congress of the Econometric Society*, 2012.

Emiliano Santoro, Ivan Petrella, Damjan Pfajfar, and Edoardo Gaffeo. Loss aversion and the asymmetric transmission of monetary policy. *Journal of Monetary Economics*, 68:19–36, 2014.

Thomas Sargent and Neil Wallace. " rational" expectations, the optimal monetary instrument, and the optimal money supply rule. *Journal of Political Economy*, 83 (2):241–54, 1975.

Thomas J Sargent. An interview with evans, george w. and seppo honkapohja. *Macroeconomic Dynamics*, 9(2005):561–583, 2005.

Ekkehart Schlicht. Isolation and aggregation in economics. Technical report, University of Munich, Department of Economics, 1985.

Ekkehart Schlicht. The moving equilibrium theorem again. *Economic Modelling*, 14 (2):271–278, 1997.

Stephanie Schmitt-Grohé and Martın Uribe. Optimal fiscal and monetary policy under sticky prices. *Journal of Economic Theory*, 114(2):198–230, 2004.

Stephanie Schmitt-Grohé and Martin Uribe. Prudential policy for peggers. Technical report, National Bureau of Economic Research, 2012a.

Stephanie Schmitt-Grohé and Martín Uribe. What's news in business cycles. *Econometrica*, 80(6):2733–2764, 2012b.

Stephanie Schmitt-Grohé and Martín Uribe. Liquidity traps: an interest-rate-based exit strategy. *The Manchester School*, 82(S1):1–14, 2014.

Stephanie Schmitt-Grohe, Martin Uribe, et al. Optimal inflation stabilization in a medium-scale macroeconomic model. *Monetary policy under inflation targeting*, pages 125–186, 2007.

Berthold Schweizer and Abe Sklar. Operations on distribution functions not derivable from operations on random variables. *Studia Mathematica*, 52(1):43–52, 1974.

R Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4(1):25–55, 1975.

Lloyd S Shapley and Martin Shubik. The assignment game i: The core. *International Journal of game theory*, 1(1):111–130, 1971.

Kevin D Sheedy. Intrinsic inflation persistence. *Journal of Monetary Economics*, 57 (8):1049–1061, 2010.

Eytan Sheshinski and Yoram Weiss. Inflation and costs of price adjustment. *Review of Economic Studies*, 44(2):287–303, 1977.

Carl P Simon and Lawrence Blume. *Mathematics for economists*, volume 7. Norton New York, 1994.

Leo Simon and Maxwell Stinchcombe. Equilibrium refinement for infinite normal-form games. *Econometrica*, 63(6):1421–43, 1995.

Robert Samuel Simon. The challenge of non-zero-sum stochastic games. *International Journal of Game Theory*, 45(1-2):191–204, 2016.

Andras Simonovits et al. Three economic applications of chebyshev's algebraic inequality. *Mathematical Social Sciences*, 30(3):207–220, 1995.

Frank Smets and Raf Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, 1(5):1123–1175, 2003.

Frank Smets and Rafael Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *The American Economic Review*, 97(3):586–606, 2007.

Peter Smith et al. Lessons from the british poll tax disaster. *National Tax Journal*, 44(4):421–436, 1991.

Richard L Smith. Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, pages 67–90, 1985.

Brian Snowdon and Howard R Vane. *Modern macroeconomics: its origins, development and current state*. Edward Elgar Publishing, 2005.

Dennis J Snower, Mewael F Tesfaselassie, et al. Job turnover, trend growth, and the long-run phillips curve. *Macroeconomic Dynamics*, 21(04):835–861, 2017.

John Stachurski. Economic dynamical systems with multiplicative noise. *Journal of Mathematical Economics*, 39(1-2):135–152, 2003.

John Stachurski. Log-linearization of stochastic economic models†. *Journal of Difference Equations and Applications*, 13(2-3):217–222, 2007.

John Stachurski, Alain Venditti, and Makoto Yano. *Nonlinear Dynamics in Equilibrium Models: Chaos, Cycles and Indeterminacy*. Springer Science & Business Media, 2012.

John Stachurski et al. Stochastic optimal growth with unbounded shock. *Journal of Economic Theory*, 106(1):40–65, 2002.

Nicholas Stern et al. The economics of climate change. *American Economic Review*, 98(2):1–37, 2008.

Joseph E Stiglitz et al. Whither socialism? *MIT Press Books*, 1, 1996.

Nancy L Stokey. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989.

Paolo Surico et al. The fed's monetary policy rule and us inflation: The case of asymmetric preferences. *Journal of Economic Dynamics and Control*, 31(1):305–324, 2007.

Lars Svensson. Inflation targeting. In *Benjamin Friedman and Michael Woodford (eds.) Handbook of Monetary Economics*, volume 3, pages 1237–1302. Elsevier, 1 edition, 2010.

Lars EO Svensson. Evaluating monetary policy. *Hoover Institution, Stanford University Book Chapters*, 2012.

Lars EO Svensson. Inflation targeting and "leaning against the wind". *International Journal of Central Banking*, 10(2):103–114, 2014.

Eric T Swanson and John C Williams. Measuring the effect of the zero lower bound on medium-and longer-term interest rates. *American Economic Review*, 104(10): 3154–3185, 2014.

Adam Szeidl. Stable invariant distribution in buffer-stock saving and stochastic growth models. *Unpublished paper.[386, 390]*, 2012.

John B Taylor. Staggered wage setting in a macro model. *American Economic Review*, 69(2):108–13, 1979.

John B Taylor. Discretion versus policy rules in practice. In *Carnegie-Rochester conference series on public policy*, volume 39, pages 195–214. Elsevier, 1993a.

John B Taylor. *Macroeconomic policy in a world economy: from econometric design to practical operation*. WW Norton, 1993b.

John B Taylor, John C Williams, et al. Simple and robust rules for monetary policy. *Handbook of Monetary Economics*, 3:829–859, 2010.

Silvana Tenreyro and Gregory Thwaites. Pushing on a string: Us monetary policy is less powerful in recessions. *American Economic Journal: Macroeconomics*, 8(4): 43–74, 2016.

Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems*, volume 140. American Mathematical Soc., 2012.

Mewael F Tesfaselassie. Trend productivity growth and the government spending multiplier. *Journal of Macroeconomics*, 37(C):197–207, 2013.

Aditi Thapar. Using private forecasts to estimate the effects of monetary policy. *Journal of Monetary Economics*, 55(4):806–824, 2008.

Peter Tillmann. Inflation targeting and regional inflation persistence: evidence from korea. *Pacific Economic Review*, 18(2):147–161, 2013.

Lorenzo Trapani. Testing for (in) finite moments. *Journal of Econometrics*, 191(1): 57–68, 2016.

Pekka Tuominen and Richard L Tweedie. Subgeometric rates of convergence of f-ergodic markov chains. *Advances in Applied Probability*, 26(03):775–798, 1994.

Kazuo Ueda. Japan's deflation and the bank of japan's experience with nontraditional monetary policy. *Journal of Money, Credit and Banking*, 44(s1):175–190, 2012.

Harald Uhlig et al. What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2): 381–419, 2005.

Martin Uribe and Stephanie Schmitt-Grohe. Pegs and pain. In *2012 Meeting Papers*, number 303. Society for Economic Dynamics, 2012.

EEC van Damme. Stability and perfection of nash equilibria. 1987.

Andrea Vaona and Guido Ascari. Regional inflation persistence: evidence from italy. *Regional Studies*, 46(4):509–523, 2012.

Hal R Varian. Microeconomic analysis. 1992.

Jesús Vázquez, Ramón María-Dolores, and Juan-Miguel Londoño. On the informational role of term structure in the us monetary policy rule. *Journal of Economic Dynamics and Control*, 37(9):1852–1871, 2013.

Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

Andreas Wagener. Chebyshev's algebraic inequality and comparative statics under uncertainty. *Mathematical Social Sciences*, 52(2):217–221, 2006.

Todd B Walker and Eric M Leeper. Information flows and news driven business cycles. *Review of Economic Dynamics*, 14(1):55–71, 2011.

Léon Walras. 1954. *Elements of pure economics*, 1874.

Peter Walters. Ergodic theory: introductory lectures. 1975.

Micah Warren. Continuum nash bargaining solutions. *arXiv preprint arXiv:1712.07202*, 2017.

Michael Weber. Nominal rigidities and asset pricing. Technical report, Society for Economic Dynamics, 2014.

Ivan Werning. Managing a liquidity trap: Monetary and fiscal policy. Technical report, National Bureau of Economic Research, Inc, 2011.

Karl Whelan. Staggered price contracts and inflation persistence: Some general results*. *International Economic Review*, 48(1):111–145, 2007.

Edmund T Whittaker. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society*, 41:63–75, 1922.

Nick Wilkinson and Matthias Klaes. *An Introduction to Behavioral Economics*. Palgrave Macmillan, 2012.

Stephen Williamson, Randall Wright, et al. New monetarist economics: models. *Handbook of Monetary Economics*, 3:25–96, 2010.

Alexander L Wolman. Sticky prices, marginal cost, and the behavior of inflation. *FRB Richmond Economic Quarterly*, 85(4):29–48, 1999.

Michael Woodford. Nonstandard indicators for monetary policy: can their usefulness be judged from forecasting regressions? In *Monetary Policy*, pages 95–115. The University of Chicago Press, 1994.

Michael Woodford. Doing without money: controlling inflation in a post-monetary world. *Review of Economic Dynamics*, 1(1):173–219, 1998.

Michael Woodford. Commentary: how should monetary policy be conducted in an era of price stability? In *Proceedings-Economic Policy Symposium-Jackson Hole*, pages 277–316. Federal Reserve Bank of Kansas City, 1999.

Michael Woodford. A neo-wicksellian framework for the analysis of monetary policy. *unpublished, Princeton University*, 2000.

Michael Woodford. The taylor rule and optimal monetary policy. *American Economic Review*, pages 232–237, 2001.

Michael Woodford. Optimal monetary stabilization policy. In *Benjamin Friedman and Michael Woodford (eds.) Handbook of Monetary Economics*, volume 3, pages 723–828. Elsevier, 1 edition, 2010.

Michael Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, 2011a.

Michael Woodford. Simple analytics of the government expenditure multiplier. *American Economic Journal: Macroeconomics*, 3(1):1–35, 2011b.

BAR-SHALOM Yaakov, Richard Osborne, Peter Willett, and Fred Daum. Crlb for likelihood functions with parameter-dependent support. *IEEE Transactions on Aerospace and Electronic Systems*, 50(3):2399–2405, 2014.

Tack Yun. Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics*, 37(2):345–370, 1996.

Tack Yun and Andrew Levin. Reconsidering the microeconomic foundations of price-setting behavior. In *2011 Meeting Papers*, number 424. Society for Economic Dynamics, 2011.

Stan Žaković, Volker Wieland, Berc Rustem, et al. Stochastic optimization and worst-case analysis in monetary policy design. *Computational Economics*, 30(4):329–347, 2007.

Andres Zambrano. The role of effort for self-insurance and its consequences for the wealth distribution. *The World Bank Economic Review*, 29(suppl_1):S118–S125, 2015.

Mark J Zbaracki, Mark Ritson, Daniel Levy, Shantanu Dutta, Mark Bergen, et al. Managerial and customer costs of price adjustment: Direct evidence from industrial markets. *The Review of Economics and Statistics*, 86(2):514–533, 2004.

Kelvin Shuangjian Zhang. Existence in multidimensional screening with general non-linear preferences. Technical report, arXiv. org, 2017.