Universitat Pompeu Fabra
Departament d'Economia i Empresa

DOCTORAT EN ECONOMIA, FINANCES I EMPRESA

ESSAYS ON INFLATION DYNAMICS AND MONETARY POLICY

Tesi Doctoral de:
Christian Höynck

Directors de la tesi:
Professor Jordi Galí i Professora Barbara Rossi

Novembre 2020

Für Marina.

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisors, Barbara Rossi, and Jordi Galí, for their support throughout this dissertation. Without their guidance and patience, this thesis would not have been possible. It has been highly rewarding to learn from their profound knowledge of economics and a gift to had their support and encouragement whenever I got stuck or lost in details. I sincerely thank you for everything.

Among the many people who helped and supported me along this way, I sincerely want to thank Lutz Weinke. He has followed and guided me at every step since the beginning. When he hired me as a student assistant in Berlin, he sparked my interest in Monetary Economics and encouraged me to follow his path to Barcelona. Thank you for your continued guidance along this way in the past years.

I also had the privilege to learn from the many other professors and the faculty at Pompeu Fabra and CREI. Among them, I would like to thank in particular Davide Debortoli for the many enriching discussions and the support he gave me along this way. Edouard Schaal and Manuel García-Santana helped me a lot during the last period of the Ph.D., both with their insightful comments, constant encouragement, and motivational speeches. I would also like to thank Isaac Baley, Christian Brownlees, Andrea Caggese, Jan Eeckhout, Geert Mesters, and the many other professors at UPF. I feel privileged to had the opportunity to learn from these people and to play some fantastic soccer games together.

I would also like to thank Marta Araque and Laura Agustí - they are keeping the house together, and without their support, I am convinced we all would not succeed in this Ph.D. journey. Marta knows the answer to any problem at UPF, and she solves them in an instant and with a great sense of humor and irony. Thank you very much for everything.

Apart from its faculty, Pompeu Fabra gave me the chance to meet wonderful people who always cheered me up and made life about so many other things than research: I would like to especially thank Adrian, Caterina, Chris, Christoph, Derrick, Dani, Donghai, Eva, Flo, Greg, Ilja, Julia, Luca, Yiru, and all those I forgot to mention and who were regular guest to my office to pick up cookies and chocolate. Sometimes, I just came to the office to talk to you or see a happy face. I am especially grateful to have met Flavio, who was always there to talk and a great coffee break companion, and Ana and Ana for our lunches, your patience with my Spanish skills, and your kindness and positivity at all times. Outside of the Ph.D. bubble, Korie was a wonderful flatmate who made Barcelona my home, far away from home. I left Barcelona, but I kept you all as friends forever. I also want to thank my dear friends Adriana, Jakob, and Karl, who always stand behind me even without being in Barcelona.

Finally, I would like to say thank you to those that are the most important to me, my family: above all, my parents, grandmother, brother, parents-in-law, and aunt-in-law. I could always rely on your encouragement, guidance, and your love. Your faith in me allowed me to embark on this journey and complete this chapter of my life.

Above all, I thank the one person whose unconditional love, support, and encouragement inspired me to pursue and complete this dissertation, my wonderful wife Marina. Words cannot describe the gratitude I have that you shared this path with me with your patience and humor. I dedicate this thesis to you.

Christian Höynck
Barcelona
November 2020

Abstract

This thesis consists of three essays that analyze the role of sectoral heterogeneity on inflation dynamics and optimal monetary policy. In the first chapter, I consider a framework where firms are connected through input-output linkages. Inflation dynamics depend on the importance of the production network to the overall economy and on the importance of particular sectors within the network. Calibrating the model to data from the United States, I document how changes to the U.S. production network can explain why the sensitivity of inflation to economic activity has declined in the past 50 years. In chapter 2, we explore the implications of market power for inflation and monetary policy. We document how the whole distribution of markups as well as the correlation between market power and price rigidity matter for the effectiveness of monetary policy and the optimal design of monetary policy. Chapter 3 provides an empirical investigation of the relative forecasting performance of core inflation for predicting underlying inflation. My results indicate that the forecast accuracy of different measures of core inflation is time-varying and that there is a trade-off in the exclusion of items between reducing noise and removing signals.

Resum

Esta tesis consta de tres artículos que analizan el rol de la heterogeneidad sectorial en las dinámicas de inflación y la política monetaria óptima. El primer capítulo parte de un marco teórico en el que las empresas están conectadas mediante vínculos "input-output". Las dinámicas de inflación dependen de la importancia de la red productiva para la economía en general, y de la importancia de sectores particulares dentro de la red. Calibrando el modelo con datos de Estados Unidos, documento cómo los cambios en la red productiva estadounidense pueden explicar por qué la sensibilidad de la inflación a la actividad económica ha disminuido en los últimos 50 años. En el segundo capítulo exploro las implicaciones del poder de mercado para la inflación y la política monetaria. Este capítulo documenta que tanto la distribución completa de márgenes de ganancia como la correlación entre poder de mercado y rigidez de precios son relevantes para la efectividad de la política monetaria y su diseño óptimo. El tercer capítulo ofrece una investigación empírica del desempeño predictivo relativo de la inflación núcleo para predecir la inflación subyacente. Mis resultados indican que la exactitud predictiva de distintas medidas de inflación núcleo varía en el tiempo, y que al excluir elementos hay un dilema entre reducir el ruido y eliminar señales.

Preface

> *My fourth question goes to the heart of monetary policy: What determines inflation? [...] I hope that researchers will strive to improve our understanding of inflation dynamics and its interactions with monetary policy.*
>
> — Janet L. Yellen, *Macroeconomic Research After the Crisis*, 2016

Inflation plays a central role in modern economics. On the one hand, it is key to the optimal conduct of monetary policy by central banks in the world. A crucial responsibility of any central bank is to control inflation. If inflation, i.e., the average rate of increase in the prices of a broad group of goods and services usually measured in terms of consumer price inflation (CPI), were not stable at a moderately low level, this would impose high costs on households and businesses. For instance, an unexpected rise in inflation tends to reduce the real purchasing power of labor income because nominal wages are generally slow or even unable ("Downward Rigidity of Nominal Wages") to adjust to price level movements. Moreover, persistently high inflation induces businesses to adjust prices more frequently than they would otherwise consider necessary. This adjustment of prices usually comes at a cost ("Menu Costs"). This vital role manifests itself by inflation control being one half of the Federal Reserve's dual mandate (besides pursuing maximum employment) and the single primary objective for the European Central Bank (ECB). Therefore, understanding what determines inflation, its relationship to monetary policy, and how to forecast inflation is of primary interest.

This thesis consists of three self-containing chapters related to the role of sectoral heterogeneity on inflation dynamics and monetary policy. Chapter one and two present and evaluate theoretical multi-sector models for analyzing production networks' role and the markup distribution. Chapter three tests for time variation in the forecasting performance of competing measures of core inflation.

Chapter one analyzes the role of changes in the structure of production networks on the flattening of the Phillips curve over the last decades. I build a multi-sector model with production networks and heterogeneity in input-output linkages and the degree of nominal rigidities. In the production network model, inflation sensitivity to the output gap depends on the topology of the economy's network. In particular, I show that two characteristics of the network matter for inflation dynamics: (i) the network multiplier and (ii) output shares. Analyzing the U.S. Input-Output structure from 1963 to 2017, I document structural changes in the production network. Calibrating the model to these sectoral changes can account for a decrease in the slope of up to 15 percent. Decomposing the aggregate effect shows that the flattening is primarily due to an increase in the centrality of sectors with more rigid prices that is incompletely reflected by compositional changes in value-added.

Chapter two, a co-authored work with Donghai Zhang, studies the role of the distribution of markups, and its changes on inflation dynamics and optimal monetary policy. The average markup of firms in the United States has increased due to the increase in the right tail of the markup distribution. We complement these empirical findings by showing that the left tail of the markup distribution has declined. We then study the implications of these findings based on a Multi-sector New Keynesian model with heterogenous markups and nominal rigidities. First, more dispersed markups lead to higher money non-neutrality in an economy with decreasing returns to scale. Second, changes in the markup distribution have minimal impact on the Phillips Multiplier – the ratio of the cumulative responses of inflation and output to a monetary policy shock – in the U.S. due to the off-setting effects of the increase in the right tail and the decrease in the markup distribution's left tail. Third, markups are negatively correlated with nominal rigidities across sectors, which has important implications for designing the optimal inflation target. Particularly, our findings challenge the conventional wisdom that the central bank should always attach a higher weight to a sector with a higher degree of nominal rigidity.

We construct the optimal inflation index and show how it has evolved over time.

The third chapter evaluates core inflation's relative performance in forecasting future medium term inflation in the U.S. I consider an approach that takes into account the possibility that the models' relative forecast performance can be time-varying. I show that the predictive ability of different measures of core inflation indeed changed dramatically over time and identify three distinct episodes. In the 1970s and until the mid-1980s, autoregressive models performed better than core inflation. From the mid-1980s until the beginning of the 2000s, all measures of core inflation outperformed headline inflation. Since the 2000s, there is no statistical difference in the predictive accuracy of both economic models. I complement these results by comparing the forecast performance of different measures of core inflation, such as permanent or temporary exclusion indices. The evidence suggests again that the relative performance is time-varying, and no measure performed best at all times. Finally, I suggest a way to test for the signaling effect of excluded components and find that missing signals from non-core inflation cannot explain why the predictive ability of core and headline are similar in recent years. These results help to understand why policymakers should monitor a wide range of core inflation indices or combine them.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1

PRODUCTION NETWORKS AND THE FLATTENING OF THE PHILLIPS CURVE

The connection between slack in the economy or level of unemployment and inflation was very strong if you go back 50 years. It has gotten weaker and weaker and weaker to the point where it's a faint heartbeat.

— Jerome Powell, *Congressional Testimony* on July 10, 2019

1.1 INTRODUCTION

The strength of the relationship between inflation and economic activity, represented by the Phillips curve, has been at the center stage of discussions among economic commentators and policymakers in the past few years. Empirical studies have found a flattening of the slope of the Phillips curve over time. This is of central importance to policymakers and central banks in particular because the sensitivity of inflation to the output gap has important implications for controlling and predicting inflation. It gives a sense of how real activity affects inflation. For instance, given a positive output gap, a smaller sensitivity implies smaller inflationary pressures. In this situation, maintaining an inflation target will become harder for a central bank. To reach the same target level, larger movements in economic activity are needed, which in turn require larger shifts in the interest rate. This is of particular concern, in times of the zero lower bound on the interest rate.

The evidence on the flattening documents that the sensitivity of inflation to output has declined by more than 50 percent, with most of the change taking place in a period after the 1980s.[1]

Understanding the sources of this shift is crucial, and economists have suggested many possible explanations. Commonly proposed explanations include the success of monetary policy in anchoring expectations (Bernanke, 2010), the credibility of the central bank (McLeay and Tenreyro, 2020), or global forces (Jorda et al., 2019). Those explanations have different implications for how optimal policy would need to change: from a larger role of fiscal policies or combined money-fiscal policies (Gali, 2020) towards rethinking inflation targeting.

In this paper, I propose a novel explanation for the flattening of the Phillips curve. I investigate the implications of changes to the production network structure of the economy for inflation dynamics. These changes go beyond changes in the value-added composition of the economy. Networks are important since firms use a variety of inputs to

[1] See for instance Ball and Mazumder (2011), Blanchard et al. (2015), Kiley (2015), Coibion and Gorodnichenko (2015), or for a recent overview Stock and Watson (2019). Studies on the wage Phillips curve include Gali and Gambetti (2019), or Hooper et al. (2020).

build their products. Thereby, they form a complex web of input-output linkages. Analysis of the input-output tables of the U.S. economy shows large changes in those interlinkages that coincide with inflation changes in the 1980s. Changes in the input-output structure have implications for the sensitivity of inflation as they alter sectoral input-output linkages. I show how the slope of the Phillips curve depends on the topology of the network. Moreover, using historical data on the input-output linkages, I find that a network-augmented Phillips curve can account for a part of the flattening of the Phillips curve since the mid-1980s.

Inflation dynamics depend on the network structure of the economy. In this paper, I study a multi-sector economy with monetary frictions in which industries are connected through input-output linkages. Additionally, I consider heterogeneity across the network structure, the degree of nominal rigidities, and markups. The first main result of this paper is that two network statistics matter for inflation dynamics: (i) *network multiplier* and (ii) *centrality captured by sectoral gross output shares*. These network statistics describe specific attributes of the input-output linkages, based on fundamentals of the economy. They have direct empirical counterparts that can be easily observed.

The network multiplier is a measure of the overall importance of the network in this economy. Total production in an economy exceeds real value-added (GDP) by intermediate good use. The network multiplier captures this excess production relative to final consumption and, therefore, how important the network channel is in an economy. The larger the network multiplier, the stronger the production networks' role in the transmission of shocks.

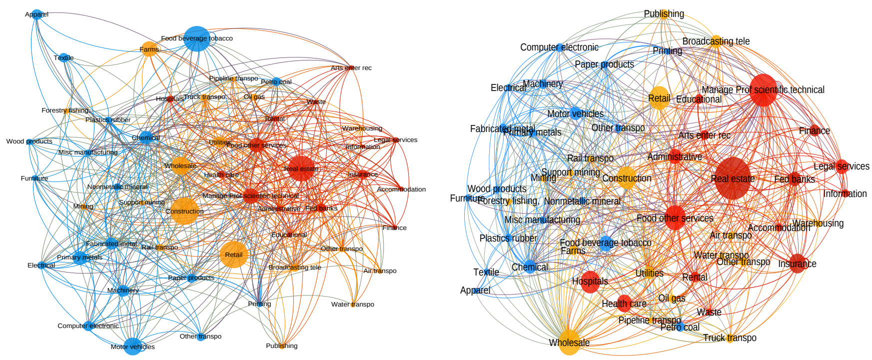
The output share is a measure of network centrality. A sector's output share captures the importance of the sector's output (i) as an input to all other sectors and (ii) for the final good. If other sectors in the economy extensively use a sector's output, its equilibrium output share will be high. Whether a sector has a high or low output share depends on the network structure. Sectors with larger output shares will contribute more to the input prices of other sectors and, therefore, to aggregate inflation dynamics. If central sectors have higher degrees of nominal rigidities, the aggregate sensitivity of inflation in this economy will be smaller.

The importance of a sector in the economy will not be given by its value added-share but rather by its gross output share. A standard multi-sector model predicts that the importance of a sectors and, therefore, the extent of their effects on aggregate inflation is related to the share of that sector in final goods aggregate demand. Instead, in the production network model, a sector can have a positive influence on the aggregate inflation dynamics even if its value-added share share is zero.

The network structure affects aggregate inflation dynamics through another channel that dampens the sensitivity of inflation: *strategic complementarities*. When the optimal price chosen by a firm depends positively (negatively) on the prices of other firms, we speak of strategic complementarities (substitutes) (see Cooper and John, 1988). Here, strategic complementarities arise because of sticky intermediate good prices. However, in my production network setting, there are two critical differences with strategic complementarities in standard formulations of intermediate goods as in Basu (1995). First, prices depend positively on the sector-specific input price instead of the aggregate price level. The sectoral input price depends on the composition of the sectoral input good, which depends on the composition of those goods constituting inputs. As an implication, the degree of strategic complementarity depends on the particular network structure of the whole economy because of those indirect supply channels. Second, the degree of strategic complementarity will be sector-specific and larger for sectors with a larger share of intermediate goods used in production.

A second implication concerns the estimation of the Phillips curve. Inflation dynamics are determined by endogenous variables in addition to the output gap. The presence of these variables biases the estimated slope coefficient of the standard Phillips curve because they are correlated with the output gap. As I demonstrate, the bias depends on the network structure. Therefore, the evolution of the Phillips curve could either be caused by a decrease in the standard slope coefficient or by a change in the bias through changes in the endogenous variables. I show that, additionally to the former effect, changes in the network structure influence the correlation between these endogenous variables and the output gap, which leads to lower estimates of the Phillips curve.

Figure 1.1: U.S. Production Network in 1963 vs. 2017



(a) Production Network in 1963 (b) Production Network in 2017

Note: Author's own calculation. This figure displays the production network corresponding to U.S. Input-Output data in 2017. Each node in the network corresponds to a sector in the 1963 input-output data, while each edge corresponds to a input-output relation between two sectors. Larger nodes represent more central sectors in terms of output shares. Color-codes represent: (i) Manufacturing (blue), (ii) Services (red) and (iii) others (orange). Source: Bureau of Economic Analysis. The figure is drawn with the software package Gephi.

The network structure of the U.S. economy has changed over time. The Bureau of Economic Analysis (BEA) provides Input-Output accounts for the U.S. economy from which a snapshot of the production network of the U.S. economy can be drawn. Panels (a) and (b) of Figure 1.1 provide network representations of the input-output linkages, in which nodes (circles) represent sectors, and edges (lines) represent input flows between sectors; the color of the node shows the originating sector. The color of nodes captures whether a sector belongs to one of three broad categories: (i) manufacturing (blue), (ii) services (red), and (iii) others (orange). Furthermore, the size of a node corresponds to the sector's centrality as measured by its output share. A thicker edge documents that the destination sector spends more expenditure on goods from the originating sector. The centrality of manufacturing firms and selected other sectors such as construction or farms, captured by the size of the blue and orange nodes, respectively, has decreased between 1963 and 2017. Conversely, the centrality of service increased as reflected by the size of red nodes in 2017. This reallocation of centrality illustrates the structural changes in the production network, while normally, structural transformation refers to the change in the value-added (GDP) shares of sectors.[2]

I study the role of the structural changes in the production network on the flattening of the Phillips curve by calibrating the multi-sector production network model to the input-output structures for each year between 1963 and 2017. I then estimate the implied sensitivity of inflation to the output gap using model-generated data for each period. The model's baseline calibration shows a flattening of the Phillips curve that is consistent with empirical evidence on the shape and timing of the flattening. While before 1980 and after 2000 the slope shows a diverging behavior in the data and the model, there is an evident flattening in the 1980s and 1990s. From the peak in the 1980s until the beginning of the 21st century, the slope of the calibrated model decreases by about 15%.

The most important channel contributing to this evolution is that changes in the production network have shifted centrality towards sectors with higher nominal rigidities: service sectors. This is equivalent to

[2]Galesi and Rachedi (2019) document an increase in the use of services as an intermediate input across all sectors (service share of intermediate inputs) and refer to this process as services deepening.

an increase in the aggregate degree of nominal rigidity in the economy. Specifically, aggregate inflation has become more rigid because service's prices are much stickier than manufacturing's prices. There is evidence from micro studies showing that service prices are more rigid than those in the manufacturing sector, e.g., Bils and Klenow (2004), Klenow and Kryvtsov (2008) or Nakamura and Steinsson (2008). Increases in the degree of nominal rigidity translate into a smaller sensitivity of inflation to the output gap in the Phillips curve.

Considering the economy's input-output structure is vital to understand the decline in the slope of the Phillips curve. Simple compositional changes in value-added fail to capture all of the explained changes to the Phillips curve. Due to sectoral reallocation, the increase in aggregate rigidity exceeds the one that would arise, considering changes in value-added shares only. Using the model, I can decompose the aggregate change to the slope estimate into the contribution of each of those two channels. I find that changes in the network structure and the value-added shares each contribute half to the explained decline in the slope of the Phillips curve.

This paper relates to the literature on sectoral heterogeneity and production networks. A growing literature studies the implications of networks on the transmission of shocks (e.g. Horvath, 2000, Acemoglu et al., 2012, Acemoglu et al., 2016 or Carvalho, 2014). In these studies, the size of the network's role in the amplification of shocks is usually related to the Leontief-Inverse matrix (Acemoglu et al., 2016 or Bigio and La'O, 2020). I contribute two new insights to this literature. First, I show that the network's impact on the transmission of shocks depends on two network statistics that capture different components of the network effects: (i) the importance of the overall network and (ii) the relative importance of sectors. Second, in the presence of nominal frictions, the network statistics and network effects become dependent on countercyclical markups. In another study, Rubbo (2020) analyses analytically optimal policy in a multi-sector framework with general input-output structures. I document the importance of the bias in estimating the slope of the Phillips curve and identify reallocation effects as the main source of the flattening in contrast to changes to the overall importance of the network.

The paper is also related to New Keynesian models with production networks. It is connected to studies that emphasize the role of networks and sectoral heterogeneity in price rigidity in amplifying the degree of aggregate monetary non-neutrality (e.g., Carvalho, 2006, Galesi and Rachedi, 2019, and Pasten et al., 2019) on government spending multipliers (Bouakez et al., 2018), or the role of price dispersion on optimal policy (Cienfuegos, 2019). My paper also ties in closely with Pasten et al. (2020), who argue that in the presence of heterogeneity in intermediate input consumption and nominal rigidities, the relevant measure of the size of a sector changes.[3] In contrast to these studies, I focus on the role of production networks (and changes to it) on inflation dynamics. Moreover, I depart by allowing for a more general network structure via heterogeneity in sectoral intermediate good shares and sectoral degrees of market power. I discuss the implications of this model for the slope of the Phillips curve and calibrate it for the U.S. economy at different points in time to compare the estimated slopes of the Phillips curve.

Section 1.2 motivates by describing the Phillips curve and reporting the problems in estimating the slope of the Phillips curve. Section 1.3 outlines the structure of the model and explains the importance of the two network statistics. Section 1.4 describes the calibration of the model and shows how the network structure, as measured by the two network statistics, has changed over time. Section 1.5 investigates inflation dynamics and predictions of the model for the sensitivity of inflation to the output gap by comparing different network economies. Section 1.6 reports the implied slope of the Phillips curve and decomposes the role of different channels. Finally, Section 1.7 concludes the paper.

[3]In particular, the effective distribution of size and centrality (out-degree) argument resembles my distinction between output shares and value-added shares.

1.2 THE PHILLIPS CURVE

At the center of macroeconomics is the theory that the economy's real and nominal side are linked through a Phillips curve relationship. Phillips (1958) provided the first formal statistical evidence on this trade-off using data on wage inflation in the U.K. Samuelson and Solow (1960) extended the "Phillips' curve" to U.S. data and price inflation. In this paper, I focus on the most widely used model of this kind, the New Keynesian Phillips curve (NKPC). It gained popularity from its theoretical microfoundations that build on early work of Fischer (1977), Taylor (1980) and Calvo (1983). It is centered around staggered price-setting by forward-looking individuals and firms.[4] The critical property of the NKPC is that inflation dynamics reflect changes in economic activity and inflationary expectations. The standard macroeconomic textbook version of the NKPC as in Woodford (2011) or Gali (2015) is

$$\pi_t = \beta E_t \pi_{t+1} + \kappa \widehat{y}_t + v_t. \tag{1.1}$$

According to this equation, inflation π_t depends on three factors: expected inflations, $E_t \pi_{t+1}$, the output gap $\widehat{y}_t$ as a measure of economic activity and v_t corresponds to cost-push shocks. Moreover, β is the time discount factor. The measure of economic activity in these models is usually marginal costs, which in turn are related to the output gap. The coefficient κ here describes the relationship between economic slack and inflationary pressures, i.e., the slope of the Phillips curve.[5]

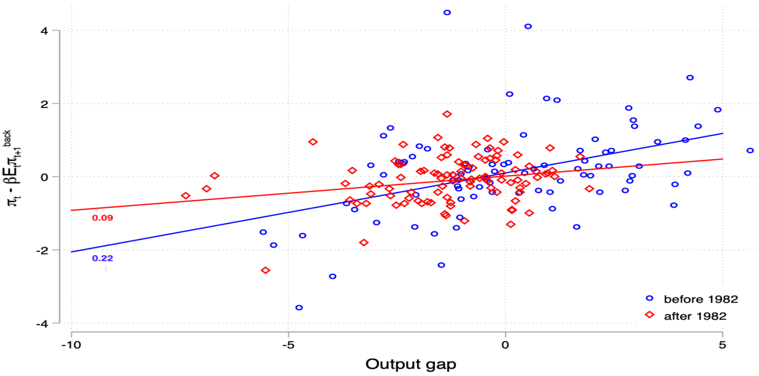
A growing literature estimates κ and reports a decrease in the coefficient over time. As an illustration of this flattening, I follow

[4] Achieved by two common ingredients: a microeconomic environment with (i) monopolistically competitive firms, and (ii) facing constraints on price-adjustment.

[5] In a standard version of this model in Gali (2015), the slope is usually given by $\kappa = (1-\theta)(1-\theta\beta)/\theta * (\sigma + \varphi)$ where θ is the Calvo parameter - the probability of not adjusting prices, β corresponds to the time discount factor, σ denotes the intertemporal elasticity of substitution and φ is the Frisch labor supply elasticity.

Stock and Watson (2019) in formalizing inflation expectations in an adaptive way, i.e., $\pi_t^e = 0.25(\pi_{t-1} + \pi_{t-2} + \pi_{t-3} + \pi_{t-4})$. This yields an accelerationist Phillips curve and the Phillips correlation instead of the slope of the NK Phillips curve, κ. I measure inflation, π_t, by year-on-year changes in PCE inflation, and the output gap, $\widehat{y}_t$, as the difference of output from its natural level, by the Congressional Budget Office's (CBO) estimates. The data is at a quarterly frequency. [6]

Figure 1.2: Changing Phillips Correlation



Note: This figure illustrates the flattening of the Phillips curve. It displays observations for predicted inflation and the output gap before and after 1985Q1 together with the implied slope estimates. Inflation is measured by the year-on-year change in PCE headline inflation. Inflation expectations (backward-looking) are captured by the four-quarter moving average of PCE inflation. The output gap is the year-on-year difference between output and the natural rate of output from the Congressional Budget Office's (CBO) estimates. Author's calculations. The figure replicates Figure 1 from Stock and Watson (2019).

The evidence of Figure 1.2 reproduces the analysis in Stock and Watson (2019) and documents that the slope of the Phillips curve was steep before 1984 (0.27) and has flattened by half since then (0.15).

[6] Details of this exercise together with more empirical evidence on the flattening can be found in the Appendix 1.8.

This flattening of the Phillips curve has led many researchers to think about possible explanations. Among those, the most prominent include anchored inflation expectations (Bernanke, 2010), the credibility of the central bank (McLeay and Tenreyro, 2020), or global forces (Jordá et al., 2019).

However, researchers face several identification problems when they seek to estimate the slope of the Phillips curve. Mavroeidis et al. (2014) report a weak identification problem that yields a wide range of estimates for κ because there is not enough variation in aggregate data. Hazell et al. (2020) or McLeay and Tenreyro (2020) attempt to overcome this problem by using regional data. Barnichon and Mesters (2020a) show that using identified demand shocks might overcome the simultaneity problem of distinguishing demand and supply shocks, the measurement error in the output gap, and unobserved inflation expectations. In this paper, I document another identification problem in estimating the Phillips curve, which arises in the presence of omitted variables. To see the relevance of omitted variables, consider that the following formula describes the Phillips curve instead of equation 1.1

$$\pi_t = \beta E_t \pi_{t+1} + \kappa \widehat{y}_t + \Psi_t + u_t, \tag{1.2}$$

where Ψ_t is an endogenous variable and $v_t = \Psi_t + u_t$ resemble the cost-push shock in equation (1.1). If the omitted variable, Ψ_t, is correlated with the output gap, $\widehat{y}_t$, the estimate of κ is biased. The bias arises because $E(\widehat{y}_t v_t) \neq 0$. I show that such a bias can arise in the presence of production networks. In the next section, I outline a multi-sector model with sectors that are related via input-output relationships. I will discuss how inflation dynamics will include an additional variable Ψ_t that depends on the structure of the production networks and introduces a bias in estimating the slope of the Phillips curve.

1.3 Model and Network Statistics

I consider a multi-sector New Keynesian Model with nominal rigidities and linkages in production via the use of sector-specific intermediate goods. In comparison to standard New Keynesian (Gali, 2015) or multi-sector models, firms use as inputs to production not only labor but also goods produced by firms from potentially all sectors of the economy. Additionally, heterogeneity in the degree of nominal rigidities, the elasticity of substitution, and intermediate good share are modeled. The model represents an extension of the standard New Keynesian model (Gali, 2015), with the sectoral models of Carvalho (2006), Carvalho and Lee (2011), Pasten et al. (2019), Cienfuegos (2019), and the intermediate good model of Basu (1995) as limiting cases.[7] The economy consists of firms, households, and a government.

The economy is composed of a continuum of firms $i \in [0,1]$ each of which belongs to a sector $k \in \{1,2,..,K\}$. Each firm produces a differentiated good that can be used either in consumption or in the production of other goods. Within each sector, firms face monopolistic competition (Dixit-Stiglitz) and produce with the same Cobb-Douglas production function that combines labor and intermediate inputs (with share γ_k). They set prices à la Calvo (1983), i.e., they can reset prices with an exogenous but sector-specific probability, θ_k.

On the consumption side, the economy is represented by a single representative household that chooses labor and aggregate consumption. The latter comprises sectoral consumption bundles, which themselves are CES aggregators of goods produced by individual firms within a sector. The labor aggregator is CES, too.

The government consists of a monetary authority, which sets the nominal interest rate following a Taylor-rule.

[7] Carvalho (2006) considers a multi-sector model with heterogeneity in nominal rigidities. Carvalho and Lee (2011) and Pasten et al. (2019) add roundabout production structures. Cienfuegos (2019) studies a model with trend inflation and without heterogeneity in the elasticity of substitution and intermediate good shares.

1.3.1 *Households*

The economy is populated by an infinitely-lived representative household with preferences on consumption, C_t, and labor, L_t. She seeks to maximize expected lifetime utility given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \frac{L_t^{1+\varphi}}{1+\varphi} \right], \tag{1.3}$$

where C_t is aggregate (final) consumption, L_t is labor input, β is the subjective time discount factor, σ is intertemporal elasticity of substitution, φ inverse of the Frisch elasticity of labor supply and E_0 is the expections operator conditional on information up to time $t = 0$.

The aggregate consumption bundle is a CD aggregator of sectoral consumption bundles

$$C_t = \prod_{k=1}^{K} C_{k,t}^{\vartheta_k}, \tag{1.4}$$

where ϑ_k is the expenditure share of sectoral consumption from sector k. Also $\sum_k \vartheta_k = 1$. The sectoral consumption bundles $C_{k,t}$ themselves are CES aggregators of the individual firms indexed in $[0, 1]$

$$C_{k,t} = \left[\int_0^1 C_{k,t}(i)^{\frac{\varepsilon_k - 1}{\varepsilon_k}} \right]^{\frac{\varepsilon_k}{\varepsilon_k - 1}}, \tag{1.5}$$

where $C_{k,t}(i)$ is the quantity of good i of sector k consumed by the household, and ε_k is the sector-specific elasticity of substitution between different goods of a sector. It is a measure of competitiveness in sectors. Note also that one usually assumes $\varepsilon_k > 1$, which implies that it is harder to substitute consumption goods from different sectors than substitute goods within the same sector.

A prominent feature of business cycle data is the sectoral comovement of output and hours worked. To capture this feature, I follow Horvath

(2000) and assume that labor provided by the household to the firms cannot move perfectly across sectors. Lee and Wolpin (2006) document that there are large mobility costs that impair the sectoral allocation of labor and Katayama and Kim (2018) document a significant degree of intersectoral labor immobility from estimates using data on sectoral hours worked. In detail, I model the aggregate labor bundle as a CES aggregator of sectoral labor supply, $L_{k,t}$, that is

$$L_t = \left[\sum_{k=1}^{K} L_{k,t}^{(1+\nu)/\nu} \right]^{\nu/(1+\nu)}, \tag{1.6}$$

where ν gives labor mobility.[8] At $\nu = \infty$, labor is perfectly mobile, and all sectors pay the same wage.

The household purchases a bundle of consumption goods and allocates the remaining income to the purchase of new bonds. She derives income from providing labor, receiving nominal profits from firms, and interest on her bond holdings. The period budget constraint is therefore given by

$$\sum_{k=1}^{K} \int_0^1 P_{k,t}(i) C_{k,t}(i) di + B_t = \sum_k W_{k,t} L_{k,t} + I_{t-1} B_t + \sum_{k=1}^{K} \int_0^1 \mathcal{D}_{k,t}(i) di, \tag{1.7}$$

for $t = 0, 1, 2, ...$, where L_t denotes the aggregate labor bundle, W_t is nominal wage, B_t represents purchases of one-period discount bonds with interest I_t, and $\sum_{k=1}^{K} \mathcal{D}_{k,t} = \sum_{k=1}^{K} \int_0^1 \Pi_{k,t}(i) di$ are aggregate dividends received from the ownership of all firms in the economy.

The household must decide on how to allocate its consumption expenditure among the different goods. The solution to this cost min-

[8] Horvath (2000) document that the idea of this specification is "to capture some degree of specificity of labor while not deviating from the representative consumer/-worker assumption".

imization of the aggregate consumption bundles yields the sectoral demand function

$$C_{k,t} = \vartheta_k \frac{P_t}{P_{k,t}} C_t, \tag{1.8}$$

where the aggregate price index is $P_t = \Pi_{k=1}^{K} \left(\frac{P_{k,t}}{\vartheta_k}\right)^{\vartheta_k}$. Similarly, cost minimization of the sectoral consumption bundles yields demand for the good of firm i in sector k

$$C_{k,t}(i) = \left(\frac{P_{k,t}(i)}{P_{k,t}}\right)^{-\varepsilon_k} C_{k,t}, \tag{1.9}$$

where sectoral price indices are $P_{k,t} = \left(\int_0^1 P_{k,t}(i)^{1-\varepsilon_k} di\right)^{\frac{1}{1-\varepsilon_k}}$ and where $P_{k,t}(i)$ denotes the price of an individual firm i in sector k. Moreover, $P_t C_t = \sum_{k=1}^{K} \int_0^1 P_{k,j,t} C_{k,j,t} dj$. Eventually, optimal allocation of sectoral labor gives labor supply

$$L_{k,t} = \left(\frac{W_{k,t}}{W_t}\right)^{\nu} L_t, \tag{1.10}$$

with $W_t = \left[(1/K) \sum_{k=1}^{K} W_{k,t}^{1+v}\right]^{1/(1+v)}$. Given the solution to the cost minimization problems, the problem of the household reduces to choosing consumption C_t, labor L_t and savings B_t to maximize lifetime utility subject to the budget constraint (1.7). The solution is described by the optimality conditions concerning labor supply and intertemporal consumption choices

$$\frac{W_t}{P_t} = L_t^{\varphi} C_t^{\sigma}, \tag{1.11}$$

$$1 = \mathrm{E}_t \left[\beta \frac{C_{t+1}^{-\sigma}}{C_t^{-\sigma}} \frac{R_t}{\Pi_{t+1}}\right], \tag{1.12}$$

where Π_{t+1} is the gross inflation rate of the aggregate price index between t and $t+1$.

1.3.2 *Firms*

Firms assemble differentiated varieties of output using labor and intermediate inputs. The goods produced are then sold as an aggregate consumption bundle to households and intermediate goods to other producers. In each sector, $k = 1, ..., K$, there is a continuum of monopolistically competitive producers indexed by $i \in [0, 1]$. Within a sector k, each firm i produces with the same Cobb-Douglas production function that combines labor input and sector-specific intermediate inputs

$$Q_{k,t}(i) = L_{k,t}(i)^{1-\gamma_k} \left(\prod_{r=1}^{K} X_{k,r,t}(i)^{\omega_{k,r}} \right)^{\gamma_k}, \tag{1.13}$$

where $Q_{k,t}(i)$ is gross output by this producer, $L_{k,t}(i)$ is labor used by firm (k, i), γ_k denotes the share of constant intermediate good use in the sector, $\omega_{k,r}$ is the relative intensity with which firms in sector k use goods produced in sector r (Input-Output shares). I assume that $\sum_{r=1}^{K} \omega_{kr} = 1 \ \forall k$. The K-by-K matrix containing the shares of intermediate input use gives the representation of the production network, denoted by $\boldsymbol{W}$. The CES - aggregator of intermediate goods purchased by firm (k, i) from all firms in sector r, $X_{k,r,t}(i)$, is given by

$$X_{k,r,t}(i) = \left(\int_0^1 X_{k,r,t}(i,j)^{\frac{\varepsilon_k - 1}{\varepsilon_k}} dj \right)^{\varepsilon_k / \varepsilon_k - 1},$$

where $X_{k,r,t}(i,j)$ is intermediate inputs purchased by firm (k, i) from firm j in sector r.

The firm's problem can then be solved in two steps. First, finding the optimal mix of inputs for a given output price that minimizes costs and, then, finding the optimal price a firm would set given these inputs.

Marginal Costs. Each firm i in sector k faces the following cost minimization problem subject to expenditure minimization of sectoral inputs, where costs are given by

$$\mathcal{C}(Q_{k,t}(i)) = \min_{L_{k,t}(i),\{X_{k,r,t}(i)\}_r} W_{k,t}L_{k,t}(i) + \sum_{r=1}^{K} P_{r,t}X_{k,r,t}(i), \tag{1.14}$$

subject to the production function (1.13). Due to the CRS technology, the cost minimization problem for firm (k, i) can be rewritten in sectoral variables only, and marginal costs are the same for all firms within the same sector. The price index for sectoral intermediate inputs is the same as for sectoral consumption goods by assuming the same elasticity of substitution in consumption and production. It yields the following formula for the nominal marginal costs of production in sector k

$$MC_{k,t} = \left(\frac{W_{k,t}}{1-\gamma_k}\right)^{1-\gamma_k} \left(\frac{P_t^k}{\gamma_k}\right)^{\gamma_k}, \tag{1.15}$$

where P_t^k is the industry-specific price index of intermediate inputs given by

$$P_t^k = \prod_{r=1}^{K} \left(\frac{P_{r,t}}{\omega_{k,r}}\right)^{\omega_{k,r}}. \tag{1.16}$$

The cost minimization has implications for firms' conditional factor demands. The firm's optimal choice of inputs, labor, and gross output, given input prices, are

$$W_{k,t}L_{k,t}(i) = (1-\gamma_k)\frac{MC_{k,t}}{P_{k,t}}P_{k,t}Q_{k,t}(i), \tag{1.17}$$

$$P_{r,t}X_{k,r,t}(i) = \gamma_k\omega_{k,r}\frac{MC_{k,t}}{P_{k,t}}P_{k,t}Q_{k,t}(i). \tag{1.18}$$

Expenditure on labor input or any particular intermediate input r is proportional to the firm's total expenditure.

Market Clearing. I can derive the total demand for goods produced by firm (i,k). This firm can either sell its product as consumption goods to the representative household, $C_{k,t}(i)$ or as intermediate input to all firms from all sectors of the economy $X_{r,k,t}(j,i)$. This implies the following market-clearing conditions

$$Q_{k,t}(i) = C_{k,t}(i) + \sum_{r=1}^{K} \int_0^1 X_{r,k,t}(j,i)dj. \tag{1.19}$$

I can use the optimality conditions from the expenditure problems to replace the CES aggregates $C_{k,t}(i) = \left(\frac{P_{k,t}(i)}{P_{k,t}}\right)^{-\varepsilon_k} C_{k,t}$ and $X_{r,k,t}(j,i) = \left(\frac{P_{k,t}(i)}{P_{k,t}}\right)^{-\varepsilon_k} X_{r,k,t}(j)$ to derive a demand for sectoral gross output

$$Q_{k,t}(i) = \left(\frac{P_{k,t}(i)}{P_{k,t}}\right)^{-\varepsilon_k} Q_{k,t}, \tag{1.20}$$

where $Q_{k,t}$ is sectoral gross output. It is defined as

$$Q_{k,t} = C_{k,t} + \sum_{r=1}^{K} X_{r,k,t}, \tag{1.21}$$

where $X_{r,k,t} = \int_0^1 X_{r,k,t}(j)dj$ is the total demand of sector r for inputs from sector k.

Price-Setting. Price-setting is modeled as in Calvo (1983), but with sector-specific probabilities as in Carvalho (2006). In particular, $(1-\theta_k)$ is the probability to reset prices in sector k. The firm's problem is then to choose the optimal price $P_{k,t}(i)$ to maximize the current market value of the profits generated while the price remains effective. Formally, firm (i, k) solves the problem

$$\max_{P_{k,t}(i)} E_t \left[\sum_{s=0}^{\infty} (\theta_k)^s \Lambda_{t,t+s} Q_{k,t+s}(i) \left[P_{k,t+s}(i) - MC_{k,t+s|s}(i) \right] \right], \quad (1.22)$$

subject to firm demand (1.20) and where $\Lambda_{t,t+s}$ is the stochastic discount factor implied by the household problem. The first-order condition is then given by

$$0 = \mathrm{E}_t \sum_{s=0}^{\infty} \Lambda_{t,t+s} \theta_k^s Q_{k,t+s}(i) \left[P^*_{k,t} - \frac{\varepsilon_k}{\varepsilon_k - 1} MC_{k,t+s} \right],$$

where $P^*_{k,t}$ is the optimal sectoral price, and $\mu_k = \varepsilon_k / (\varepsilon_k - 1)$ is the sectoral markup absent nominal rigidities. Thus, firms resetting their prices will choose a price that equals the markup over their current and expected marginal costs. The weights depend on the economy's discount rate and the probability of the firm's price remaining unset until each respective horizon.

Defining relative prices as $p^*_{k,t} = \frac{P^*_{k,t}}{P_t}$ and rewriting the optimality condition in the standard recursive form, yields

$$p^*_{k,t} = \frac{\varepsilon_k}{\varepsilon_k - 1} \frac{\psi_{k,t}}{\Delta_{k,t}}, \quad (1.23)$$

where $\psi_{k,t}$ and $\Delta_{k,t}$ are auxiliary variables that represent expected discounted values of marginal costs and revenues. They are defined recursively as

$$\psi_{k,t} = Q_{k,t} C_t^{-\sigma} mc_{k,t} + \theta_k \beta \mathrm{E}_t \left[\Pi_{k,t+1}^{\varepsilon_k} \psi_{k,t+1} \right], \quad (1.24)$$

$$\Delta_{k,t} = Q_{k,t} C_t^{-\sigma} + \theta_k \beta \mathrm{E}_t \left[\frac{\Pi_{k,t+1}^{\varepsilon_k}}{\Pi_{t+1}} \Delta_{k,t+1} \right], \tag{1.25}$$

where $mc_{k,t} = MC_{k,t}/P_t$ are real marginal costs and $\Pi_{k,t-1,t+s} = P_{k,t+s}/P_{k,t-1}$ is the gross nominal inflation rate.

The Calvo environment implies that sectoral price dynamics are described by

$$P_{k,t} = \left[(1-\theta_k) \left(P_{k,t}^{*} \right)^{1-\varepsilon_k} + \theta_k \left(P_{k,t-1} \right)^{1-\varepsilon_k} \right]^{1/1-\varepsilon_k}. \tag{1.26}$$

1.3.3 *Monetary Policy*

The monetary authority sets the short-term nominal interest rate, I_t, according to the following Taylor rule with value-added output and aggregate inflation

$$\frac{I_t}{\bar{I}} = \left(\frac{Y_t}{\bar{Y}} \right)^{\phi_c} \left(\frac{\Pi_t}{\bar{\Pi}} \right)^{\phi_\pi} e^{\upsilon_t}, \tag{1.27}$$

where $\Pi_t = \Pi_{k=1}^{K} \Pi_{k,t}^{\vartheta_k}$ is the aggregate inflation rate, variables with a bar denote steady-state values, Y_t is aggregate nominal value-added and υ_t is a monetary policy shock that follows an AR(1) process. In this model, the respective real measure is the value-added output, Y_t, instead of gross output, Q_t, and the inflation index relevant for household consumption is Π_t. The coefficients ϕ_c and ϕ_π measure the degree to which the monetary authority adjusts the nominal interest rate in response to changes in the consumption-based inflation rate and changes in the value-added output, respectively.

1.3.4 *Equilibrium*

Before turning to the model's log-linearized solution and the Phillips curve, in this section, I describe the model's equilibrium system. In particular, I stress its properties related to the determinants of marginal costs and the connection between gross output and value-added output. The equilibrium is described by the firms' and household's optimality conditions along with market clearing conditions.

Aggregation and Value-Added Output. In this economy, real aggregate value-added (i.e. real GDP), Y_t, is equal to consumption, C_t. Let $\mathcal{Y}_{k,t}(i)$ denote the nominal value-added of producer i in sector k. It is defined as the value of gross output produced by this firm abstracting the value of intermediate inputs it is using, i.e.

$$\mathcal{Y}_{k,t}(i) = P_{k,t}(i)Q_{k,t}(i) - \sum_{r=1}^{K} P_{r,t}X_{k,r,t}(i). \tag{1.28}$$

Aggregating over all real value-added output of all producers in sector k

$$\mathcal{Y}_{k,t} = \int_0^1 \mathcal{Y}_{k,t}(i)di = P_{k,t}Q_{k,t} - \sum_{r=1}^{K} P_{r,t}X_{k,r,t}, \tag{1.29}$$

where $X_{k,r,t} = \int_0^1 X_{k,r,t}(i)di$ by intermediate input clearing condition.

I can aggregate nominal dividends by using

$$\mathcal{D}_{k,t} = \int_0^1 D_{k,t}(i)di = P_{k,t}Q_{k,t} - W_{k,t}L_{k,t} - \sum_{r=1}^{K} P_{r,t}X_{k,r,t} = \mathcal{Y}_{k,t} - W_{k,t}L_{k,t}, \tag{1.30}$$

where I use the labor market clearing condition $L_{k,t} = \int_0^1 L_{k,t}(i)di$.

Eventually, substituting into the household's budget constraint (1.7), aggregate dividends and the bond market-clearing $B_t = 0$, I obtain

$$P_t C_t = \sum_{k=1}^{K} W_{k,t} L_{k,t} + \sum_{k=1}^{K} \mathcal{D}_{k,t} = \sum_{k=1}^{K} \mathcal{Y}_{k,t}. \tag{1.31}$$

The aggregate nominal value-added equals the nominal value of total consumption. Real aggregate value-added (i.e., real GDP), Y_t, can then be derived by deflating the nominal aggregate value-added by the aggregate consumption price index, P_t, that is

$$Y_t = \frac{\sum_{k=1}^{K} \mathcal{Y}_{k,t}}{P_t} = C_t. \tag{1.32}$$

Wages and Total Expenditure. The role of total production in affecting marginal costs becomes clearer when deriving the equilibrium wages in this economy. Combining the labor market clearing condition $L_{k,t} = \int_0^1 L_{k,t}(i) di$ with labor demand from firms (1.17) yields sectoral labor demand

$$L_{k,t} = (1 - \gamma_k) \frac{mc_{k,t}}{p_{k,t}} p_{k,t} Q_{k,t} d_{k,t} w_{k,t}^{-1}, \tag{1.33}$$

where $mc_{k,t} = \frac{MC_{k,t}}{P_t}$, $p_{k,t} = \frac{P_{k,t}}{P_t}$, and $w_{k,t} = \frac{W_{k,t}}{P_t}$ are real sectoral marginal costs, relative sectoral prices and real sectoral wages respectively. Moreover, $d_{k,t} = \int_0^1 \left(\frac{P_{k,t}(i)}{P_{k,t}} \right)^{-\varepsilon_k} di$ is within sector price dispersion. This corresponds to price dispersion in a one-sector model, and as shown in Gali (2015), around a zero inflation steady state, price dispersion is approximately zero. Thus, for expositional purposes, I will not carry it along in the derivations that follow since it becomes negligable up to a first order approximation.[9] The exposition shows that

[9] It will become relevant under the assumptions of positive trend inflation (Ascari and Sbordone, 2014). Then this will introduce the propagation of sectoral price dispersions as discussed in Cienfuegos (2019).

sectoral labor demand depends on the total sectoral expenditure of the sector. This expenditure is the share of total revenue from production that is not spent on markups. Labor demand is increasing in the real value of sectoral production and decreasing in sectoral markups.

One can combine labor supply (1.11) and labor demand (1.17) with marginal costs (1.15) to solve for real wages

$$w_{k,t} = \frac{W_{k,t}}{P_t} = \left(\left(\frac{1-\gamma_k}{\gamma_k} \right)^{\gamma_k} (p_t^k)^{\gamma_k} Q_{k,t} \right)^{\frac{\varphi}{1+\gamma_k \varphi}} C_t^{\frac{\sigma}{1+\gamma_k \varphi}} \tag{1.34}$$

The wage rate depends on firms' demand for labor input and the household's labor supply via the following mechanisms. Wages increase in firms' demand for labor input if their total production increases, $Q_{k,t}$, and in the cost of intermediate goods, p_t^k, as they can substitute labor inputs for intermediate inputs. On the other hand, wages increase in household's labor supply if their demand for the aggregate consumption good, C_t, increases, or the disutility from working, φ, falls.

Marginal costs. Replacing wages (1.34) in marginal costs (1.15), we can show that the average marginal cost in sector k yields

$$mc_{k,t} = \phi_k Y_t^{\frac{(\sigma+\varphi)(1-\gamma_k)}{1+\gamma_k \varphi}} \left(\frac{p_{k,t} Q_{k,t}}{Y_t} \right)^{\frac{\varphi(1-\gamma_k)}{1+\gamma_k \varphi}} \left(\frac{p_t^{k\,\gamma_k}}{p_{k,t}} \right)^{\frac{(1+\varphi)}{1+\gamma_k \varphi}}, \tag{1.35}$$

where $\phi_k = \frac{1}{1-\gamma_k} \left(\frac{1-\gamma_k}{\gamma_k} \right)^{\frac{\gamma_k(1+\varphi)}{1+\gamma_k \varphi}}$ is a constant.

In this economy, marginal costs are affected by three components: (i) the aggregate demand channel, Y_t, (ii) the real value of sectoral output, and (iii) the price of intermediate inputs. While the first channel is standard, the other two are due to production networks. This is, however, only a partial equilibrium analysis since the three variables are endogenous.

Linking Sectoral Production to GDP. The next step in the derivation is to solve for the real value of sectoral production, $p_{k,t}Q_{k,t}$, in terms of real value-added, Y_t. In particular, I will show that sectoral markups will affect other industries through the production network channel. Therefore, I will introduce two network statistics that measure (i) the share of production from a sector in total production (output share) and (ii) the share of intermediate goods in total production (network multiplier). I use the definitions of real value-added, the market-clearing condition (1.15) and the budget constraint (1.7) to obtain the following characterization for the real value of sectoral production in terms of real aggregate value-added, i.e., GDP.

Proposition 1.3.1 *Let $Q_t \equiv \sum_r p_{r,t}Q_{r,t}$ denote the real value of total production (gross output), $\delta_{k,t} \equiv p_{k,t}Q_{k,t}/Q_t$ be the output share of sector k, and $\Phi_t^{NM} \equiv Q_t/Y_t$ the network multiplier of the economy. The real value of sectoral production in the multi-sector economy with production networks and nominal frictions is linked to aggregate real value-added and given by*

$$p_{k,t}Q_{k,t} = \delta_{k,t}\Phi_t^{NM}Y_t, \tag{1.36}$$

with

$$\begin{aligned}
\Phi_t^{NM}\left(\overrightarrow{\frac{1}{\mathcal{M}_t}}, \delta_t\right) &= \left[1 - \mathbf{1}'(\gamma \odot \overrightarrow{\frac{1}{\mathcal{M}_t}} \odot \delta_t)\right]^{-1}, \\
\delta_t\left(\overrightarrow{\frac{1}{\mathcal{M}_t}}\right) &= \left[\mathcal{I} - (W' - V_C\mathbf{1}')\left(\mathbf{1}\left(\gamma \odot \overrightarrow{\frac{1}{\mathcal{M}_t}}\right)'\right)\right]^{-1} V_C,
\end{aligned} \tag{1.37}$$

where $\odot$ denotes the Hadamard (entrywise) product and $\overrightarrow{\frac{1}{\mathcal{M}_t}}$ the $Kx1$ vector of sectoral real marginal cost deflated by sectoral prices which is also related to the inverse of sectoral markups $\frac{1}{\mathcal{M}_{k,t}} = \frac{mc_{k,t}}{p_{k,t}}$. Also, γ,

V_C and δ_t are $Kx1$ vectors of sectoral intermediate good shares, sectoral consumption shares and sectoral output shares, respectively. W' is the inverse of the input-output matrix, reflecting how much intermediate goods each sector k provides to all other sectors.

Proposition 1.3.1 states that the real value of sectoral production is proportional to real value-added and depends on the topology of the production network. The network structure is captured by two statistics, (i) the network multiplier of the economy and (ii) the output share of a sector. Those two statistics – and hence the relationship between sectoral production and value-added – are affected by variations of markups across sectors.

In the next two sections, I will provide insight into the two network statistics and explain why we can think about the network multiplier as a measure of the importance of the network to the economy and the output share as a measure of the centrality of a sector.

Network Multiplier. The network multiplier, Φ_t^{NM}, provides a link between real value-added production, Y_t, and the real value of total production (gross output), $Q_t \equiv \sum_r p_{r,t} Q_{r,t}$, in the economy. In a model without intermediate goods, the multiplier would be one. In the multi-sector model with production networks, to produce one more unit of the aggregate consumption good, additional production units are produced that will be used as intermediate goods. Therefore, the network multiplier will be larger than one.

The larger the network multiplier, the more labor is needed to produce the same consumption unit. One can think about this as a proxy for the length of the production chain in this economy. The longer the production chain, the more intermediate goods, and hence labor input is needed for production.

To derive the network multiplier, I rewrite the budget constraint (1.7) in terms of real value-added and the total value of production

$$Q_t = Y_t + \sum_{k=1}^{K} \gamma_k \frac{mc_{k,t}}{p_{k,t}} \delta_{k,t} Q_t, \tag{1.38}$$

where I use the definition of the output shares $p_{k,t}Q_{k,t} = \delta_{k,t}Q_t$. The real value of total production exceeds real aggregate value-added by the aggregate expenditure on intermediate goods.

This recursive expression of the real total value of production can be iterated to obtain a representation capturing all direct and indirect expenditure effects along with the production network

$$Q_t = Y_t + \underbrace{\sum_{k=1}^{K} \gamma_k \frac{mc_{k,t}}{p_{k,t}} \delta_{k,t} Y_t + \sum_{k=1}^{K} \gamma_k \frac{mc_{k,t}}{p_{k,t}} \delta_{k,t} \sum_{r=1}^{K} \gamma_r \frac{mc_{k,t}}{p_{k,t}} \delta_{r,t} Y_t + \dots}_{\text{network component}}. \tag{1.39}$$

The first term captures the household's expenditure on the consumption good. The second term captures the indirect expenditure on intermediate goods as a proportion of total consumption. The third and higher terms reflect the expenditure on intermediate goods higher in the production chain. All of those indirect effects decay at the rate given by the intermediate good share, γ_k, and the markup, $\frac{1}{\mathcal{M}_(k,t)}$.[10] The size of the decay is weighted by the importance of a particular sector, reflected by its output share, $\delta_{k,t}$. The network component captures these indirect effects. It increases if large (in terms of output shares) sectors have larger intermediate good shares and smaller markups. Here we can see how distortions in markups can propagate through the network and

[10]If we assume the intermediate share to be constant, $\gamma_k = \gamma$, we would find

$$Q_t = Y_t + \gamma \sum_{k=1}^{K} \frac{mc_{k,t}}{p_{k,t}} \delta_{k,t} Y_t + \gamma^2 \sum_{k=1}^{K} \frac{mc_{k,t}}{p_{k,t}} \delta_{k,t} \sum_{r=1}^{K} \frac{mc_{k,t}}{p_{k,t}} \delta_{r,t} Y_t + \dots,$$

i.e., higher-order terms are devaluated by the rate γ.

change total expenditure. In the New Keynesian model, markups are countercyclical in response to demand shocks, which means that the network component and the network multiplier increase in booms.

The decay of the network multiplier has another intriguing interpretation. The smaller the decay, the longer will be the production chain. If the network structure is such that more central sectors have larger intermediate shares, then the total decay of the network will be smaller, and the total multiplier larger. I will investigate this mechanism further when I look at different examples of networks.

Rewriting the sums in vector form, I can rewrite the last expression as

$$Q_t = \underbrace{\left[1 - \mathbf{1}' \left(\gamma \odot \overrightarrow{\frac{1}{\mathcal{M}_t}} \odot \delta_t\right)\right]^{-1}}_{=\Phi_t^{NM}} Y_t, \tag{1.40}$$

where γ, $\overrightarrow{\frac{1}{\mathcal{M}_t}}$ and δ_t are $Kx1$ vectors of sectoral intermediate good shares, sectoral markups and sectoral output shares, and Q_t as well as Y_t are scalars.

The network multiplier, Φ_t^{NM}, shows how much total production is needed in order for the household to consume one unit of the consumption good. By capturing the size of the network component relative to the direct expenditure on the consumption good, it is, therefore, a measure of the relative importance of the network in this economy. As long as intermediate good shares are positive, the network component will be non-zero, and the network multiplier will be larger than one. The network multiplier is not constant as it depends on markups and the output shares. In the next section, I develop a closed-form expression for output shares.

Output Shares. The output share, $\delta_{k,t}$, provides a link between the real value of sectoral production and the real value of total production.

The output share will be equal to the consumption share in a model without intermediate goods, V_C. However, in the multi-sector model with production networks, each sector also provides intermediate goods to other sectors. From the system of market-clearing conditions (1.21) in combination with sectoral intermediate good demands (1.18), the demand for sectoral production yields

$$p_{k,t}Q_{k,t} = \vartheta_k C_t + \sum_{r=1}^{K} \omega_{r,k}\gamma_r \frac{mc_{r,t}}{p_{r,t}} p_{r,t}Q_{r,t}, \tag{1.41}$$

where the first part on the right-hand side represents the household's direct demand for goods from sector k. This demand is fully described by the consumption share of goods from sector k in the total demand of the household for the aggregate consumption good, ϑ_k, i.e., it reflects the preferences of the household.[11] The second component is the demand from other sectors r that use sector k's good as an intermediate input. It is given by the share sector r spends on goods from sector k, $\omega_{r,k}$, relative to its total intermediate good expenditure.

Combining the previous expression with the budget constraint (1.7) and dividing by the total real value of production yields an iterated representation of the output shares

$$\delta_{k,t} = \vartheta_k + \underbrace{\sum_{r=1}^{K} \widetilde{\omega}_{r,k}\vartheta_r + \sum_{r=1}^{K}\sum_{s=1}^{K} \widetilde{\omega}_{s,r}\widetilde{\omega}_{r,k}\vartheta_s + \ldots,}_{\text{network component}} \tag{1.42}$$

where $\widetilde{\omega}_{r,k,t} = (\omega_{r,k} - \vartheta_k)\gamma_r \frac{1}{\mathcal{M}_{r,t}}$ is a weighting matrix.

The first term represents the direct demand for goods from sector k from the household. The first term of the network component captures the importance of sector k to its immediate customers, firms that are

[11] By the Cobb-Douglas assumption on the consumption aggregator, this is thus a constant fraction of total consumption.

directly connected to k. The specific contribution of sector r to k's output share depends on sector r's own share, $\delta_{r,t}$, and on a weighting matrix, $\widetilde{\omega}_{r,k}$. This weighting matrix depends on the network weight connecting both sectors, $\omega_{r,k}$ and the intermediate good expenditure of sector r. The second term of the network component captures the indirect importance of k through sectors that buy inputs from sector k's customers. In other words, this is the indirect demand from the customers of the customers of sector k. The third and higher-order terms capture the importance of k through customers that are one or more further steps away from sector k.

Again, we can rewrite the last equation in vector form to represent the relationship between the vector of sectoral real values of production, $\overrightarrow{PQ}_t$, and the total real value of production, Q_t, in this economy

$$\overrightarrow{PQ}_t = \underbrace{(\mathcal{I} - \widetilde{W}_t)^{-1} V_C}_{=\delta_t} Q_t, \tag{1.43}$$

where $\widetilde{W}_t = (W' - V_C \mathbf{1}')(\mathbf{1}(\gamma \odot \overrightarrow{\tfrac{1}{\mathcal{M}_t}})')$.

The output share summarizes the network structure by specifying how much sectoral production $p_{k,t}Q_{k,t}$ is needed in order to satisfy a given demand for gross output, $Q_t = \sum_r p_{r,t}Q_{r,t}$. In fact, output shares are equal to the *network centrality* of Katz (1953), i.e. they capture the relative importance of each node (sector) in a network (aggregate economy). A sector is important if its outdegree is larger than its consumption share, i.e. $(W'_k - V_{C,k} > 0)$, and if its customers have large intermediate good shares, γ_k, or small markups, $\mathcal{M}_{k,t}$. In the absence of intermediate goods or markups, the output share equals the consumption share, V_C. Again, fluctuations in markups in other sectors are transmitted through the network via intermediate good use, W.

1.3.5 *Steady State and Network Examples*

In this section, I highlight some key features of network statistics in steady-state. Therefore, I examine some examples of network structures to provide insight into the network statistics introduced before. Moreover, I discuss how both statistics can be directly calculated from the data.

The main difference between a zero-inflation steady-state and the equilibrium is that markups are constant and given by $\bar{\mathcal{M}} = \frac{\varepsilon_k}{\varepsilon_k - 1}$ in the former. Evaluating network statistics and marginal costs in steady-state yields

$$\bar{\Phi}^{NM} = \left[1 - \mathbf{1}'(\gamma \odot \overrightarrow{\frac{1}{\bar{\mathcal{M}}}} \odot \delta)\right]^{-1}, \tag{1.44}$$

$$\bar{\delta} = \left[\mathcal{I} - (W' - V_C \mathbf{1}')\left(\mathbf{1}\left(\gamma \odot \overrightarrow{\frac{1}{\bar{\mathcal{M}}}}\right)'\right)\right]^{-1} V_C. \tag{1.45}$$

This expression shows that the two network statistics only depend on network characteristics, the sectoral markups, and are independent of nominal rigidities. In the dynamic model (i.e., outside the steady-state), the markups will vary and, hence, both network statistics.

Another interesting implication comes from these equations. Given observables for intermediate good shares, value-added shares, the input-output tables, and markups, we can calculate both network characteristics directly (without additional assumptions) from the data. They have direct empirical counterparts in the data and can be observed at the yearly frequency from input-output tables. Thus, they could potentially be used as sufficient statistics for the slope of the PC.

Before I turn to the calibration part, where I show how the network statistics have changed over time, I will outline how different features of hypothetical network economies affect the two network statistics.

Example 1: Change in Network Multiplier. This economy has two sectors. I assume that the household preference weights on each good are the same, $V_C = [0.5; 0.5]$. Markups are symmetric and equal to 20%, i.e. $\frac{1}{\mathbb{M}} = 5/6$. Moreover, I consider two networks that differ in their intermediate shares

$$W_1 = W_2 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \text{ and } \gamma_1 = \begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix} \text{ and } \gamma_2 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

Both networks feature a symmetric roundabout production network: Sectors 1 and 2 equally spend their input expenditure on inputs from Sectors 1 and 2. The equivalence of consumption shares and outdegrees implies that both sectors have the same output share in steady-state ($\delta = V_C$), with each sector having 50% of the market. However, Network 2 has a smaller network multiplier of 1.71 because of its lower intermediate share than Network 1 of 2.67.

In this example, the network multiplier changes without changes in output shares. This will become important later to decompose changes in the slope through the lens of a symmetric network model.

Example 2: A (non-)"irrelevant" sector. Keeping the markup structure of Example 1, and adjusting the houshold preferences such that Sector 1 becomes irrelevant for households $V_C = [0; 1]$, Network 3 and is given by

$$W_3 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \text{ and } \gamma_3 = \begin{pmatrix} 0.75 \\ 0.75 \end{pmatrix}.$$

This network represents a star network: Sectors 1 and 2 spend all of their input expenditure on inputs from Sectors 1. Therefore, Sector one that is irrelevant from the household perspective is the central sector of this economy from the network perspective: Sector 1 has non-zero output shares as $\delta = [0.63; 0.37] \neq V_C$. By the symmetry of the intermediate good shares and markups, Network 3 has the same network multiplier as Network 1, 2.67.

Hence, consumption weights do no longer characterize the importance of a sector for the economy. This illustrates that Hulten's law (Hulten, 1978) does not hold in this economy. The impact of a sectoral TFP shock is not equal to the sector's share in total value-added.

1.3.6 *Equilibrium Conditions and Dynamics*

The equilibrium is described by a system of $7N + 3$ equations to pin down the $7N + 3$ endogenous variables: sectoral variables $\{\Pi_{k,t}$, $p_{k,t}$, $Q_{k,t}$, $mc_{k,t}$, $d_{k,t}$, $\psi_{k,t}, \Delta_{k,t}\}_{k=1}^{K}$ and aggregate $\{Y_t, I_t, \Pi_t\}$ variables given the monetary shocks e^{υ_t}. The three aggregate equations are the Euler equation (1.12), the Taylor rule (1.27) and the aggregate labor supply (1.11). The sectoral equations are output demand (1.21), labor demand and marginal costs (1.15) those that determine the optimal pricing decision (1.23), (1.24), and (1.25).

1.4 Calibration and Network Changes

This section describes the baseline calibration of the model and the data sources. One of the objectives of the calibration is to compute the model's implied slope of the Phillips curve over time. I will allow for time-variation in the calibration of different parameters: the production network, i.e. (i) the composition of sectoral intermediate goods, W, which will be derived from the Input-Output tables, as well as (ii) sectoral intermediate good shares, γ_k, and (iii) the size of each sector as measured by its value-added share, ϑ_k.

In the second part, I outline how the production network in the U.S. has structurally changed in the past decades as represented by the two network statistics introduced in the last section: (i) network multiplier and (ii) output shares. In this respect, I show that services have not only become more central in value-added terms but also with respect to the network structure. It follows a discussion of examples and characteristics that describe those sectors that have become most central.

1.4.1 *Calibration*

Starting from the sectoral definitions of the “summary level” of the Input-Output tables from the BEA, I excluded the government sector to be consistent with the model. Moreover, the specification of sectors has changed from 1996 to 1997. To account for these changes in the classification, I merged five sectors to be consistent with the 1963 specification. Eventually, the dataset covers 53 sectors at roughly the 3-digit NAICS level from 1963 to 2017.

Production Network. I use data from the Bureau of Economic Analysis (BEA) on the flow of goods from each industry in the U.S. economy to other industries. The aggregated industries defined by the BEA sum to gross domestic product and therefore cover the entire economy. The Input-output tables are available at an annual frequency and show the dollar value of goods produced, for example, in industry i that industry j uses as inputs. For each sector, I use this information to derive the composition of intermediate goods $\omega_{k,r}$, final demand $\vartheta_{,k}$ as well as the intermediate goods share, γ_k.

Frequency of Price Changes. The frequency of price changes is calibrated using data from Pasten et al. (2019).[12] They calculate monthly frequencies using confidential microdata underlying the Bureau of Labor Statistic’s (BLS) Producer Price Index (PPI). Based on these frequencies of price adjustments at the goods level, they aggregate these into the

[12] I am grateful to Michael Weber for sharing this data.

350-sector industry-level definitions of the Bureau of Economic Analysis (BEA). To map them into the 53-sector specification, I compute the median frequency within each 3-digit sector. The monthly frequencies are transformed to match the quarterly calibration of the model.

Other Parameters. I calibrate the model at the quarterly frequency using standard parameter values in the literature. The discount factor is assumed to be $\beta = 0.99$, which implies an annual steady-state return on financial assets of about 4 percent. It is also assumed a unitary intertemporal elasticity of substitution and inverse Frisch elasticity of labor supply as well as labor mobility $\sigma = \nu = \varphi = 1$. As to the interest rate rule coefficients, it is assumed $\phi_\pi = 1.5$ and $\phi_c = 0.5/4$. The persistence parameter of monetary shocks is $\rho_m = 0.5$. Finally, the constant elasticity of substitution is set equal to $\varepsilon = 6$ in order to match a steady-state markup of 20%.[13] Table 1.1 summarizes the calibration of the other business cycle parameters. Table 1.1 summarizes the calibrated values for all parameters.

Parameter	Description	Value
σ	Constant relative risk aversion	1
ν	Inverse of Frisch elasticity of labor	1
ε	Constant elasticity of substitution	6
ϕ_π	Inflationary response of the Taylor Rule	1.5
ϕ_c	Output-gap response of the Taylor Rule	0.5/4

Table 1.1: Calibration Homogenous Parameters

[13] In the empirical part of the paper, I abstract from heterogeneity in the elasticity of substitution in the benchmark case.

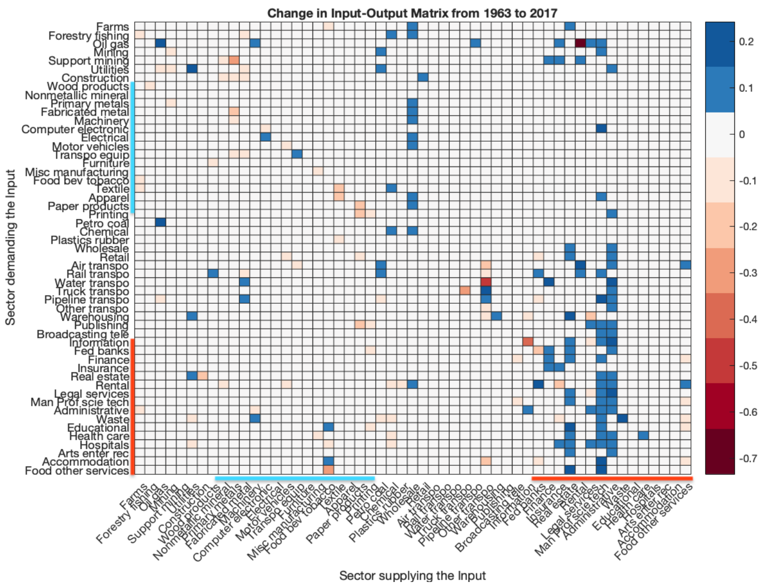
1.4.2 *Structural Change in the Production Network*

This section documents how the input-output network structure of the U.S. economy has changed over time. I use the BEA input-output data to analyze the U.S. economy over a long time span (from 1963 to 2017) at an annual frequency. Importantly, I find that a sectoral change from manufacturing to services did not only take place for final demand but also in terms of the network structure. Consistent with a process of service deepening (Galesi and Rachedi, 2019), I find that certain, usually services-based industries have become more important in terms of centrality or intermediate good provision in the network over time. In the second part of the section, I study how these changes are reflected in the two network statistics from Section 1.3.

Changes in the Input-Output Structure over Time. First, I will look at changes to the economy's network structure by comparing the input-output tables from 1963 to 2017. An input-output matrix shows the contribution from all the other sectors to the intermediate goods of sector k (vertical axis). Figure 1.3 displays the difference in Input-output matrices as created by the BEA handbook guidelines. Increases in the use of an input from sector j (horizontal axis) by sector k (vertical axis) from 1963 to 2017 are colored in blue, while decreases in the use are colored in red. Over time, there is a substantial increase in the use of services – input use from services are on the right of the horizontal axis – across the board. Notably, two service sectors besides FIRE have become significantly more important, as illustrated by two blue vertical lines: (i) management of companies and enterprises and professional, scientific, and technical services ("Man prof scie tech") and (ii) administrative services.

Which Industries have Become More or Less central? In this section, I re-confirm the previous result by analyzing how the input-output relationships have changed over time using our first network

Figure 1.3: U.S. Input-Output Structure Over Time

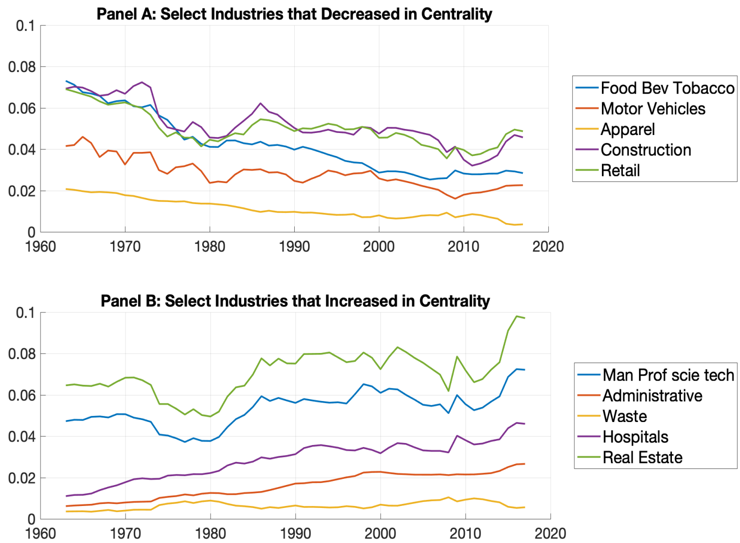


Note: This figure shows the change in the Input-output matrix from 1963 to 2017 as created by the BEA handbook guidelines. While increases in the intensity of demands or provisions are colored blue, decreases in the importance of an edge are colored red.

characteristic: *output shares*. The output share, δ_k, is a centrality measure, as outlined in Section 1.3. Centrality is one way of measuring each node's (industry) relative importance in a graph (network). It is particularly important because centrality considers an industry's connections to other industries and the strength of these connections, and how connected the other industries are. In this way, an industry will tend to have a high centrality measure if it is connected to other industries with high centrality.

Figure 1.4 shows the evolution of centrality in terms of output shares of selected industries. Panel A of Figure 1.4 reports those sectors that decreased in centrality. These were mostly manufacturing firms such as "food, beverage and tobacco products", or "motor vehicles, and

FIGURE 1.4: Centrality of Selected Industries Over Time



Note: This figure shows the change in the output shares, $\delta_{k,t}$ from 1963 to 2017 for selected industries. Panel A shows industries that were among the top decreases in centrality, while Panel B displays industries that increased most in centrality. Sources: Bureau of Economic Analysis and author's own calculations.

parts". They were among the most central firms in the 1960s but have declined dramatically in their importance. We can find 4 Manufacturing firms among the least central industries in 2017, while in 1963, none was manufacturing. In 2017, half of the most central industries were services. This is confirmed by Panel B of 1.4 which shows industries that increased in centrality. Important contributors to this change is the increasing importance of the health sector with "hospitals and nursing and residential care facilities" or "ambulatory health care services" as well as the FIRE sector represented by "real estate" and "insurance carriers and related activities".

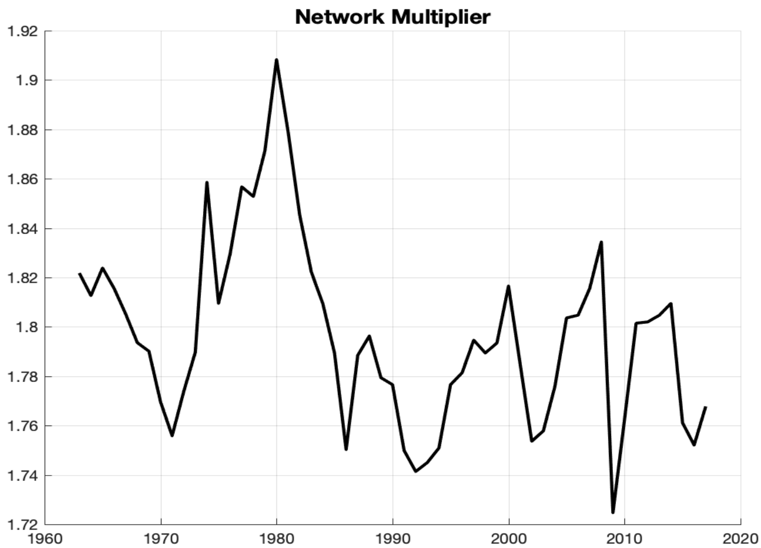
Does this Resemble the Structural Transformation in Value-Added Shares? Industries with high centrality do not necessarily also have high value-added. For instance, "management, professional, scientific, and technical services" is among the top rank increases in terms of centrality; however, the industry's GDP share has stayed almost unchanged. The same can be observed for sectors that became less central. While "primary metals", "machinery" and "fabricated metal products" are among the industries that became less central, their GDP share did not change. Centrality measures an industry's importance as part of the input-output network, not necessarily its importance within GDP. GDP counts only the amount of goods and services that go into final uses, e.g., consumption or investment (value-added). GDP excludes input-output flows since intermediate goods are excluded from the value-added output. Therefore, industries that are large providers of intermediates goods but not final goods can have high centrality but low value-added, and vice versa. Across all industries and all years, the correlation between an industry's value-added and its centrality is 0.41, a moderate level that suggests a not-very-strong link between the two measures.

Therefore, industries that are important from a network perspective may not necessarily be important from a GDP perspective. Conversely, just because an industry is important in terms of value-added, does not imply that it has a large role in the input-output network.

No Trend in Network Multiplier Over Time. This section documents how the second network statistic has changed over time: The *network multiplier*. The network multiplier, Φ_t^{NM}, is a measure of the importance of the network to the economy. It takes values above one to reflect that to produce one more unit of the aggregate final good, additional units of intermediate goods need to be produced. The network multiplier's empirical counterpart from the BEA input-output tables is the ratio of total gross production to value-added output. Figure 1.5

shows the evolution of the U.S. network multiplier. There was a spike at the end of the 1970s, but overall there is no trend in the importance of the production network. On average, the network multiplier is 1.8, which translates into an excess production of 80% due to the use of intermediate goods.

Figure 1.5: Network Multiplier Over Time



Note: This figure shows the change in the network multiplier over time. The empirical counterpart in the data is the ratio of gross production to value-added production. Sources: Bureau of Economic Analysis and author's own calculations.

In summary, this section has documented important changes to the U.S. network over time. While there was no change in the overall importance of networks to the economy – represented by a flat network multiplier – there has been an increase in the importance of service sectors to the U.S. production network. There was not only a reallocation between sectors in terms of value-added shares (usually referred to as *structural transformation*), but also in terms of the centrality of sectors (*structural change in production networks*). In particular, services became the most central sector in the economy (service deepening). This has important implications on the sensitivity of inflation to the output gap, as discussed in the next section.

1.5 INFLATION DYNAMICS

In this section, I discuss how the production network is reflected in the Phillips curve. I will show how the multi-sector model with production networks compares to other models concerning inflation's sensitivity to the output gap. Before analyzing the impact of the structural changes in the network statistics identified in the previous section, I will derive the sensitivity in different cases to study the contribution of each feature of the model on inflation dynamics.

To derive the Phillips curve in this environment, I start by log-linearizing the model's equilibrium conditions around a zero-inflation steady-state and analyzing the resulting system of difference equations. Unless otherwise noted, I use the $\wedge$ symbol on top of a variable to indicate the deviation from its steady-state value.[14]

1.5.1 *The Sectoral Phillips Curves*

As standard in New Keynesian models, inflation dynamics are described by a forward-looking relationship between inflation and marginal costs

$$\widehat{\pi}_{k,t} = \beta E_t \widehat{\pi}_{k,t+1} + \kappa_k (\widehat{mc}_{k,t} - \widehat{p}_{k,t}), \tag{1.46}$$

where $\kappa_k = (1-\theta_k)(1-\beta\theta_k)/\theta_k$. Due to heterogeneities in the multi-sector model, each sector will be described by a separate Phillips curve.

[14] One could also express the solution of the Phillips curve in terms of deviations from the natural level of output as common in the NK literature. In our model, without productivity shocks, the natural level of output is a constant, and the slope coefficient is unchanged. For illustrative purposes and a better comparison to other multi-sector models, I deviate from this convention.

In the next step, I combine the previous expression with markups from equation (1.35) to describe sectoral Phillips curves by

$$\widehat{\pi}_{k,t} = \beta \mathrm{E}_t \widehat{\pi}_{k,t+1} + \frac{(1-\gamma_k)}{1+\gamma_k\varphi} \Phi_k^{Std} \widehat{y}_t + \Psi_{k,t}, \tag{1.47}$$

where $\Phi_k^{Std} = \kappa_k(\sigma+\varphi)$ is the slope coefficient in the standard one-sector NK model (with sector-specific θ_k) and $\Psi_{k,t}$ is an endogenous term

$$\Psi_{k,t} = \underbrace{\kappa_k \frac{\varphi(1-\gamma_k)}{1+\gamma_k\varphi}(\widehat{\delta}_{k,t} + \widehat{\Phi}_t^{NM})}_{\Psi_{k,t}^{NW}} + \underbrace{\kappa_k \frac{(1+\varphi)}{1+\gamma_k\varphi}(\gamma_k \widehat{p}_t^k - \widehat{p}_{k,t})}_{\Psi_{k,t}^{SC}}, \tag{1.48}$$

where I define $\Psi_{k,t}^{NW}$ as the *network component* and $\Psi_{k,t}^{SC}$ as the *strategic complementarity component.*

The expression shows that sectoral inflation dynamics in this economy depend on three variables additional to future expected inflation: (i) the output gap as in the standard Phillips curve, (ii) network changes, either due to the network multiplier or to the sectoral output share, $(\widehat{\Phi}_t^{NM} + \widehat{\delta}_{k,t})$ and (iii) sectoral input price and relative price gaps, $\gamma_k \widehat{p}_t^k - \widehat{p}_{k,t}$. The latter two variables are two channels that determine inflation but are ignored in one-sector models.

The production network enters in three ways. First, the coefficients in front of the three variables depend on the sectoral intermediate share, γ_k. With more intensive intermediate good use, the first two channels (via the wage channel) become less important, and marginal costs depend more on input price gaps, $\widehat{p}_t^k$. Second, these relative input price gaps depend on the input-output network via W since they are defined by $\widehat{p}_t^k = \sum_r \omega_{k,r} \widehat{p}_{r,t}$. Three, the network component depends on (i) the network multiplier and (ii) output shares as delineated in Section 1.3.

There are numerous implications from the sectoral Phillips curves in (1.47). First, relative price gaps in the sectoral Phillips curves add persistence into the inflation dynamics (see Woodford, 2011). In particular, relative prices are lagged endogenous variables, and, thus, they introduce a backward-looking component in the determination of inflation. This occurs due to the multi-sector structure with heterogeneity in nominal rigidity as in Woodford (2011), which is represented by $\widehat{p}_{k,t}$. Additionally, the presence of intermediate goods in production introduces relative price gaps via input price gaps. Second, due to the presence of relative price gaps and sectoral production in the determination of aggregate inflation, there will not be a so-called "divine coincidence" (Blanchard and Gali, 2007). Instead, the central bank will face a trade-off between the stabilization of inflation and the output gap.

1.5.2 *Aggregate Inflation Dynamics and the Slope of the Phillips Curve*

Using the definition of the aggregate price index, $\widehat{\pi}_t = \sum_k \vartheta_k \widehat{\pi}_{k,t}$, the aggregate Phillips curve is a weighted average of the sectoral Phillips curves given by

$$\widehat{\pi}_t = \beta \mathrm{E}_t \widehat{\pi}_{t+1} + \Phi^{Std} \widehat{y}_t + \Psi_t, \tag{1.49}$$

where $\Phi^{Std} = \sum_k \vartheta_k \kappa_k \frac{(1-\gamma_k)}{1+\gamma_k \varphi}(\sigma + \varphi)$ and $\Psi_t = \vartheta_k \Psi_{k,t}$.

Aggregate inflation dynamics in this economy are determined by the sum of sectoral dynamics of the output gap, and inflation expectations, $\beta \mathrm{E}_t \widehat{\pi}_{t+1}$. Moreover, they depend on an additional endogenous variable, Ψ_t, that is the sum of a *network component*, $\Psi_t^{NW} = \sum_k \vartheta_k \Psi_t^{NW}$, and a *strategic complementarity component*, $\Psi_t^{SC} = \sum_k \vartheta_k \Psi_{k,t}^{SC}$.

What is the slope of the Phillips curve, i.e., the sensitivity of aggregate inflation to the output gap? Naturally, one would say it is the coefficient in front of the output gap, Φ^{Std}. However, in the multi-sector model with production networks, this answer is incomplete because the endogenous variable is correlated with the output gap, i.e., $E[\widehat{y}_t \Psi_t] \neq 0$. Thus, the slope of the Phillips curve is the sum of the standard slope, Φ^{Std}, and the sensitivity of the endogenous variable, Ψ_t, with respect to the output gap. To find the slope of the Phillips curve, the correlation between the output gap and the endogenous variable needs to be calculated. If this correlation is negative, then the presence of the production network has a dampening effect on the slope of the Phillips curve.

What would an econometrician, estimating the sensitivity of inflation to the output gap uncover? Estimating the model in equation (1.1), while the data-generating model is described by equation (1.49), he would get a biased estimate of κ. The reason is that we can think of the endogenous variable, Ψ_t , as cost-push shocks, v_t, in equation (1.1). This generates an omitted variable bias because $E[\widehat{y}_t v_t] \neq 0$, and introduces a bias in estimating the slope of the standard Phillips curve. The flattening of the Phillips curve could then either be because (i) a decline in the standard slope, Φ^{Std}, or (ii) a change in the bias, $E[\widehat{y}_t \Psi_t]$. The bias depends on the *network component* and the *strategic complementarity component*. In the next section, I will quantify the size of the bias in a calibrated multi-sector model.

A few additional things are noteworthy about the *strategic complementarity component*. It shows the presence of strategic complementarities in price-setting. When the optimal price chosen by a firm depends positively (negatively) on the prices of other firms, we speak of strategic complementarities (substitutes) (Cooper and John, 1988)[15]. There are different sources for strategic complementarities such as kinked demand curves (Kimball, 1995) or factor attachments (see Basu, 2005). In the present case, strategic complementarities arise because of (i) relative sectoral demand, $p_{k,t}$, and (ii) sticky intermediate good prices, $\gamma_k p_t^k$. However, in the production network setting, there are two distinct differences to strategic complementarities from general standard formulations of intermediate goods (Basu, 2005).

First, prices depend positively on the sector-specific input price, p_t^k, instead of the aggregate price level, p_t. As an implication, the degree

[15]In the simple setting of CJ, strategic complementarity arises if the profit of a firm $V(p_i, p_{-i})$ is such that $V_{12}(p_i, p_{-i}) > 0$.

of strategic complementarity will be sector-specific and depend on the share of intermediate goods used in production. Second, due to the network structure, also indirect effects will be present. The price-setting of a firm will depend on the prices of their suppliers. However, since those use intermediate goods in production, their prices depend on the suppliers' suppliers' prices and so on. Consequently, the degree of strategic complementarity depends on the particular network structure of the whole economy. In particular, the interaction of suppliers' stickiness, θ_k, and intermediate good share, γ_k, will impact downstream sectors; e.g., if an important supplier is rigid, its downstream sectors will exhibit large degrees of strategic complementarity, too. This feature reduces the sensitivity of real marginal cost and hence aggregate inflation to changes in aggregate demand (real rigidity, Ball and Romer, 1990).

Before calculating the size of the bias in the slope of the Phillips curve in different models, in the next section, I will look at one particular case of the economy from Section 1.3 and illustrate the implications for the Phillips curve and its slope.

Multi-Sector Model without intermediate goods. Here, I will derive the Phillips curve of a multi-sector model without intermediate goods, $\gamma_k = 0$, but with heterogeneity in nominal rigidities (Woodford, 2011 Ch. 3 or Carvalho, 2006). The aggregate inflation dynamics in this model are given by

$$\widehat{\pi}_t = \mathrm{E}_t \widehat{\pi}_{t+1} + \Phi^{Std} \widehat{y}_t - (1+\varphi) \sum_k \vartheta_k \kappa_k \widehat{p}_{k,t} \tag{1.50}$$

where $\Phi^{Std} = \sum_k \vartheta_k \kappa_k (\sigma + \varphi)$ is the average slope coefficient.

Here, the standard slope of the Phillips curve is given by Φ^{Std}. The last term in Equation (1.50) is due to heterogeneity in nominal rigidities.[16] The endogenous variable will be based only on the relative price,

[16] In the absence of heterogeneity in nominal rigidities, prices are symmetric: $\sum_k \vartheta_k \widehat{p}_{k,t} = 0$.

$\widehat{p}_{k,t}$, and not the input price part, $\gamma_k \widehat{p}_t^k$, of the strategic complementary component, Ψ_t^{SC}. The network component is absent because, without intermediate goods, gross production is equivalent to final demand. The econometrician estimating equation 1.1, will again recover a biased estimate of Φ^{Std} due to the omitted variable bias, $\mathrm{E}[\widehat{y}_t \Psi_t^{SC}] \neq 0$. For this reason, oil prices or prices of imported goods are commonly added to the estimation of standard Phillips curves to capture the effect of relative price changes, Ψ_t^{SC}, and recover the standard slope, Φ^{Std}. However, this will not be enough in the presence of production networks since one also needs to account for the network component, Ψ_t^{NW}.

1.5.3 *Implications for the Slope in Different Economies*

In this section, to quantify the importance of the estimation bias, I calibrate the economy to the input-output structure of 2007. I will compute the model implied slope, which I define as the slope an econometrician will estimate using (1.1) and compare it to the standard slope coefficient, Φ^{Std}, and the bias from the multi-sector model, $\mathrm{E}[\widehat{y}_t \Psi_t]$. Moreover, I will provide a decomposition of the bias into the network component, Ψ_t^{NW}, and the strategic complementarity component, Ψ_t^{SC}. Before I explain how I calculate the model implied slope, I will describe different versions of the fully calibrated model that I will use to compare the contributions of different elements of the model for the slope and the bias.

Different Economies. In particular, I start with the homogeneous multi-sector model (Case 1) and will add step-wise (i) heterogeneity in the frequency of price change (Case 2), (ii) in the intermediate share (Case 4), and (iii) asymmetry in the network structure (Case 5). Eventually, Case 6 represents the fully calibrated U.S. economy in 2007.

Case 1: "Homogeneous Multi-Sector Economy": Multi-sector model, no production network, homogeneous frequency of price adjustment

Case 2: "Heterogenous Multi-Sector Economy": Multi-sector model, no production network, heterogenous frequency of price adjustment

Case 3: "Symmetric Input-Output Economy": Symmetric production network, homogeneous intermediate share, homogeneous frequency of price adjustment

Case 4: "Heterogeneous Intermediate Shares Economy": Symmetric production network, heterogenous intermediate shares, homogeneous frequency of price adjustment

Case 5: "Asymmetric Network Economy": Asymmetric production network, homogeneous intermediate shares, homogeneous frequency of price adjustment

Case 6: "Full 2007 Economy": Asymmetric production network, heterogenous intermediate shares, heterogenous frequency of price adjustment

Monte Carlo Evidence. This section will calculate the model implied slope of the Phillips curve in the following steps. In the first step, the system formed by the equilibrium equations described in Section 1.3.6 and calibrated to the U.S. economy in 2007 as in Section 1.4.1 is simulated 2000 times for 200 periods, respectively. The simulated data series of aggregate inflation, $\widehat{\pi}_t$, the output gap, $\widehat{y}_t$, the network component, Ψ_t^{NW}, and the strategic complementarity component, Ψ_t^{SC}, are collected at each repetition and will depend on the specific series of monetary policy shocks in the considered Taylor rule, (1.27).

In the second step, I use the simulated data to calculate the model implied slope and the biases. The standard slope coefficient, Φ^{Std}, can be calculated directly using the parameters of the model and equation (1.49).

The model implied slope, is the sensitivity of inflation to changes in the output gap an econometrician would estimate from model (1.1) and the simulated data, i.e., the estimate $\widehat{\kappa}$ in $\pi_t^{sim} = \beta E_t \pi_{t+1}^{sim} + \kappa \widehat{y}_t^{sim} + v_t$. To calculate the size of the bias from the network component, $Bias^{NW} = \mathrm{E}[\widehat{y}_t \Psi_t^{NW}]$, I will project the network component, $\Psi_t^{NW,sim}$, on the output gap, $\widehat{y}_t^{sim}$. The bias from the strategic complementarity component, $Bias^{SC}$, is calculated in the same way. Finally, I will compare the sum of the standard slope coefficient, and the biases with the model implied slope, $\widehat{\kappa}$.

Table 1.2 reports the mean estimates of the Phillips curve components, averaged over 2000 repetitions. Using medians, instead, does not change the results. The first column of Table 1.2 corresponds to the multi-sector economy without production networks or intermediate goods and with a homogeneous frequency of price adjustment. In this economy, the slope of the Phillips curve is equal to the standard slope coefficient, Φ^{Std}, and the econometrician is estimating exactly this coefficient without a bias, $\widehat{\kappa}$.[17]

The multi-sector economy without intermediate goods but with heterogeneous degrees of nominal rigidities is illustrated in Column two. The standard slope coefficient, Φ^{Std}, increases in comparison to Case 1 because of a concave relationship between sectoral price rigidities and the sectoral slope coefficient (see Imbs et al., 2011, or Chapter 2 for more details). Heterogeneous price rigidity also introduces relative price gaps as in Carlstrom et al. (2006). As outlined in equation (1.50), relative price gaps are an additional endogenous variable correlated with the output gap. From the simulated data, we can see that the bias in estimating the Phillips curve, $Bias^{SC}$, is negative. Relative price gaps are increasing in response to the output gap, but the coefficient in front of relative price gaps, $\widehat{p}_{k,t}$, is negative in equation (1.49). However, the sensitivity of inflation to the output gap is given by the sum of the

[17]Specifically, the slope collapses to the textbook one-sector economy, $\Phi^{Std} = \kappa(\bar{\theta})(\sigma + \varphi)$, where $\bar{\theta}$ is the average degree of price rigidity in the economy.

Table 1.2: The Slope Components of the Phillips Curve in Multi-Sector Models

	Multi-Sector		Production Network			2007 Economy
	Homogenous Frequency (Case 1)	Heterogenous Frequency (Case 2)	Symmetric Network (Case 3)	Heterogenous Int. Shares (Case 4)	Asymmetric Network (Case 5)	(Case 6)
			Panel (a) Slope Components			
Φ^{Std}	0.48	2.63	0.94	0.83	0.94	0.83
$Bias^{NW}$	0	0	0.33	0.28	0.25	0.22
$Bias^{SC}$	0	-2.21	-1.04	-0.88	-0.94	-0.81
$\sum$	0.48	0.42	0.23	0.225	0.246	0.249
			Panel (b) Model Implied Slope			
$\widehat{\kappa}$	0.51	0.448	0.243	0.237	0.26	0.263

Note: This table compares the slope components (Panel (a)) and the model implied slope (Panel (b)) from different calibrated economies of the multi-sector model, calculated using simulated data. The standard slope coefficient, Φ^{Std}, is calculated directly from the parameters. The biases $Bias^{NW}$ (or $Bias^{SC}$) are derived by projecting simulated data for the network (or strategic complementarities) component of the endogenous variable, $\Psi_t^{NW,sim}$, on simulated data for the output gap. Panel (b) shows the result of projecting simulated data for inflation on simulated data for the output gap. The first two columns represent multi-sector models without intermediate goods and homogenous or heterogeneous degrees of nominal rigidity respectively. Columns 3 and 5 consider the production network model with homogenous intermediate shares but a symmetric network or the actual (asymmetric) network. Column 4 adds heterogeneity in intermediate good shares to the symmetric network in column 3. Column 6 represents the full production network model calibrated to 2007.

standard slope and the bias. Overall, the bias has a dampening effect, and the implied slope of the Phillips curve, $\widehat{\kappa}$, is smaller than in Case 1. The smaller slope increases monetary non-neutrality in this type of model as in Carvalho (2006), or Carvalho and Schwartzman (2015), where amplification depends on the distribution of price rigidity across sectors.

The production network is introduced step-wise in the next four columns. The network structure affects the different components of the aggregate Phillips curve in three ways. First, the size of the standard slope coefficient, Φ^{Std}, is reduced. The reason is that wages are less elastic to changes in the output gap in the presence of intermediate goods. The relationship is proportional to the sectoral share of intermediate good use, γ_k. In particular, the dampening is stronger for heterogenous intermediate good shares (Case 4), indicating that sectors with large intermediate good shares tend to have larger degrees of nominal rigidity in the U.S. economy.

Second, the bias from strategic complementarities decreases. The strategic complementarities component is a combination of input price gaps and relative price gaps, $\gamma_k \widehat{p}_t^k - \widehat{p}_{k,t}$. The bias from intermediate good price gaps is smaller and of opposite sign and reduces the overall bias from strategic complementarities. The dampening effect is decreasing in the sectoral share of intermediate good use, γ_k.

Third, the network bias is positive because the output gap and the network component, Ψ_t^{NW}, are positively related. This is consistent with Bils et al. (2018) who document that the intermediate good use is pro-cyclical in the U.S. economy. The size of the bias depends on the network multiplier, $\widehat{\Phi}_t^{NM}$, and the centrality, $\widehat{\delta}_{k,t}$, of sticky sectors. Case 5 illustrates that in the asymmetric network, the centrality of sticky sectors has increased, which decreases the network bias relative to the symmetric network in Case 3.

The sum of the biases and the standard slope coefficient is again equivalent to the model implied slope. However, a small difference arises from a bias due to inflation expectations.[18]

In summary, the biases and the model implied slope depend on the structure of the production network. The analysis of this section has shown two main results: First, in the U.S. production network, the standard slope coefficient, and the bias from strategic complementarities cancel each other out. The elasticity of inflation to the output gap is almost completely related to the bias from the network component. Second, comparing a symmetric and an asymmetric production network shows that the bias of the network component decreases if the centrality of sticky sectors in the economy increases.

1.6 The Slope of the Phillips Curve over Time

In this section, I study the evolution of the Phillips curve through the lens of the multi-sector model. The flattening of the Phillips curve could either be because of (i) a decline in the standard slope or (ii) a change in the bias. To answer this question, I combine the model with historical data on input-output linkages for the U.S. economy from 1963 until 2017.[19] The identification strategy is that changes in the network structure will be reflected as changes to firms' technology in the model. Precisely, I match parameters of the production function in the model to changes in the expenditure shares from the input-output tables in the data. The exercise is then to fit the model to the production structure at each point in time, simulate data from the model, and estimate the model implied slope of the Phillips curve from equation (1.1), $\widehat{\kappa}$. This approach allows us (i) to decompose the changes in the model implied

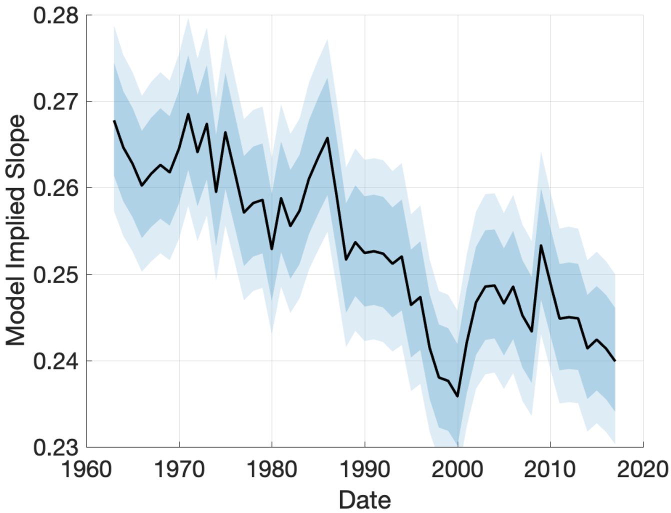
[18] In the calculations of the slope components, we did not use inflation expectations, while we assumed the econometrician observes them for the implied slope. In the appendix, I calculate the size of the inflation expectations bias.

[19] An alternative way is to directly estimate equation (1.49) in the data. Due to data limitations related to historical sectoral input prices, I follow the structural approach in this paper.

slope into variations in the standard slope coefficient and the biases, and (ii) to perform counterfactual exercises.

1.6.1 *The Production Network Model Implied Slope*

FIGURE 1.6: The Model Implied Slope of the Phillips Curve



Note: This figure reports the model implied slope of the Phillips curve over time. The slope is constructed by simulating time series for inflation and the output gap from the multi-sector model at different points in time and then regressing inflation on the output gap. The shaded regions report point-wise 68% and 90% credible sets.

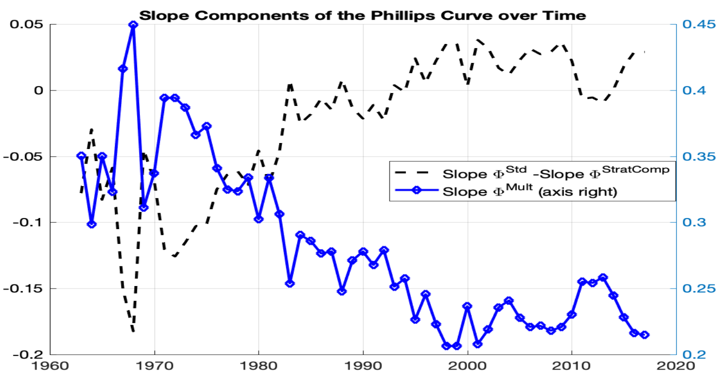
The model implied slope of the Phillips curve, $\widehat{\kappa}$, is flattening over time. The solid black line in Figure 1.6 depicts the model implied slope of the Phillips curve together with confidence bands derived from estimates of different repetitions.[20] The line matches the shape and the timing of the identified changes to the Phillips curve. Until the

[20] Compared to the evidence from the data, the level of the slope is larger than the data would suggest. A possible explanation includes an aggregation effect that is prominent in multi-sector models as compared to the slope in a one-sector model. In the following analysis, I focus on the change over time.

mid-1980s, the slope is relatively flat. From the mid-1980s until the beginning of the 2000s, the slope is flattening, and since then, we can see a slight rebound and a generally diverging behavior of the slope. From the peak in the 1980s until the beginning of the 2000s, the slope decreases by about 15%. This corresponds to 25 to 50 percent of the total decrease in the sensitivity of inflation to the output gap that was estimated in the literature (see Stock and Watson, 2019 and an analysis in the Appendix 1.8).

1.6.2 *The Role of the Network Bias*

Figure 1.7: Decomposition of the Model Implied Slope



Note: This figure displays the evolution of the the bias from the network component, $Bias^{SC}$, (blue solid line) together with the sum of the standard slope of the Phillips curve, Φ^{Std}, and the bias from strategic complementarities, $Bias^{SC}$, (black dashed line).

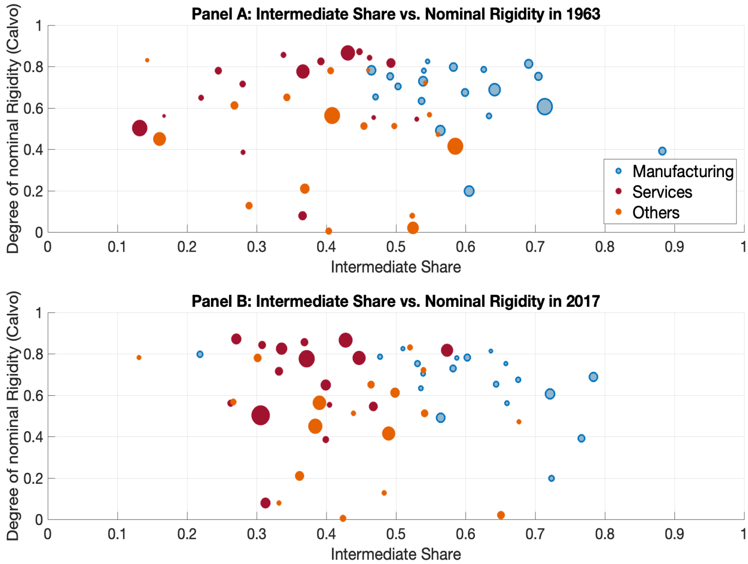
The evidence from the previous section documented that changes to the production network can explain a significant part of the flattening of the Phillips curve. To understand the sources of this flattening, Figure 1.7 decomposes the evolution of the model implied slope into the different components introduced in the previous section. The solid blue line depicts the evolution of the bias that arises from the network component, Ψ_t^{NW}. The network bias falls by about 50% over time and, therefore, constitutes the main source of the flattening of the slope. In contrast, the other two components of the Phillips curve, i.e., the standard slope coefficient and the bias from strategic complementarities, cannot explain the flattening. Figure 1.7 illustrates that the sum of these two components (black dashed line) is increasing over time. In recent years, the sum of them is close to zero. In conclusion, the change in the network bias is the most important source of the flattening of the Phillips curve. The next two sections analyze the roles of structural changes to the two network statistics in the decrease in the bias: (i) the role of the centrality of sectors and (ii) the role of the network multiplier.

1.6.3 *Channel 1: The Role of the Reallocation of Centrality to Sticky Sectors.*

Section 1.4.2 documented that the production structure of the U.S. economy has changed over time. In particular, the centrality of services as measured by their output shares, $\delta_{k,t}$ increased. These changes have an important effect on the sensitivity of inflation to economic activity because sectors have different price-setting behavior. There is ample evidence from micro studies showing that prices are more rigid in the service sector (e.g., Bils and Klenow, 2004, Klenow and Kryvtsov, 2008, or Nakamura and Steinsson, 2008).

Sectors that have become more central have more rigid prices. To give a first impression of this reallocation, Figure 1.8 compares the sectoral heterogeneity in nominal rigidities in 1963 and 2017. Each sector is represented by a bubble, where the size of the bubble corresponds to the centrality of the sector measured by its output share, $\delta_{k,t}$, in 1963 and 2017, respectively. The figure shows that service sectors (red bubbles) increase in centrality and that they display a larger rigidity. In contrast, a number of formerly important manufacturing sectors (blue) become

Figure 1.8: Relationship Between Centrality, Price Rigidity and Intermediate Share



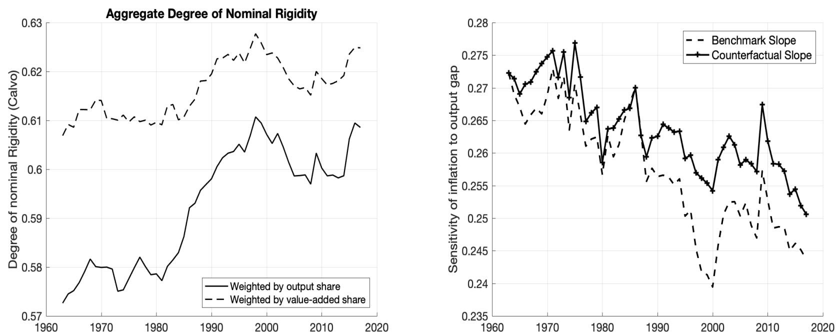
Note: This figure shows the sectoral relationship between the sectoral degree of nominal rigidity (vertical axis), intermediate good share (horizontal axis) and centrality of a sector (bubble size). Panels A and B show the relationship in 1963 and 2017.

irrelevant in 2017.

The degree of nominal rigidity plays an important role in New Keynesian models. When the probability of adjusting prices decreases, the average duration of firm's prices increases, so it becomes more likely that those prices are away from their optimal level. Sectors with larger nominal rigidities (larger Calvo parameters, θ_k), therefore, have a smaller sensitivity of inflation to the output gap.

Changes to the network have increased the aggregate degree of nominal rigidity by shifting production towards stickier sectors - i.e., sectors with a larger Calvo parameter. Panel (a) of Figure 1.9 displays the average degree of nominal rigidity in the economy, where the weights

FIGURE 1.9: The Model Implied Evolution of the Phillips Multiplier in the U.S.: The Role of Sectoral Heterogeneity in Price Rigidity



(a) Aggregate Degree of Nominal Rigidity (b) Counterfactual: Homogeneous Price Rigidity

Note: Panel (a) of this Figure displays the aggregate (weighted average) degree of nominal rigidity in the U.S. economy, where the weights are sectoral output shares, $\delta_{k,t}$, (black solid line) or value-added shares, $\vartheta_{k,t}$, (black dashed line). Panel (b) shows the evolution of the model implied slope of the Phillips (black dashed line) together with a counterfactual slope that is derived from a multi-sector model with homogeneous price rigidity across sectors.

are either the value-added shares (dashed line) or output shares (dashed line) of the respective sector. The average degree of nominal rigidity increases over time, particularly, after 1980 irrespective of the weighting. Inspecting the scales, however, the change is bigger if sectors are weighted by their output shares.[21]

To formally test for the role of heterogeneity in the degree of nominal rigidity, I perform a counterfactual exercise. Panel (b) of Figure 1.9 shows the path of the model implied slope of the Phillips curve (black

[21] Although the scale in the change of the degree of nominal rigidity appears small, the impact on the slope of the Phillips curve is significant, due to the non-linear relationship. The increase of rigidity from 0.57 to 0.61 in the graph, decreases the slope by over 20% from 0.33 to 0.25, considering $\kappa = (1-\theta) * (1-\beta\theta)/\theta$.

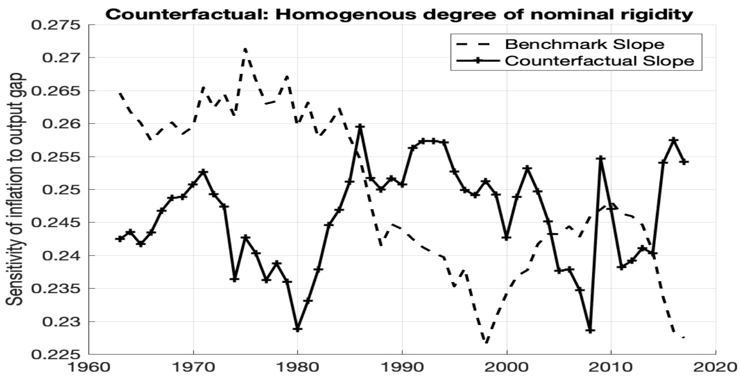
dashed line) and compares it to a counterfactual slope (solid black line) that is derived from a multi-sector model with a homogeneous degree of nominal rigidities across sectors. In the latter economy, each sector's nominal rigidity is set to the average rigidity in the economy at that point in time. Figure 1.9 confirms the importance of the reallocation channel because preventing this channel; the slope of the Phillips does not flatten over time.

The Role of Structural Transformation. The increase in aggregate rigidity could arise from changes in the value-added shares of sectors (structural transformation) or changes in output shares. The changes in the sectoral GDP shares as measured by final demand from the BEA Use table show a decline in the size of manufacturing and an increase in the size of services. The increase is strongest for services related to health care.[22] Nevertheless, Panel (a) of Figure 1.9 documents that the impact of those changes on aggregate price rigidity is smaller than those from variation in output shares.

To disentangle the role of structural transformation on the slope, Figure 1.10 reports the result of a counterfactual exercise, where the value-added shares, V_C, are kept constant at their levels from 1963 (solid black line). Any change in the solid line is due to structural changes in the network structure instead of changes in the value-added shares. Figure 1.10 shows that structural transformation in networks and value-added shares each contribute half to the total decline in the slope (black dashed line). To understand the flattening of the Phillips curve, it is important to consider both (i) structural transformation in value-added shares and (ii) changes to the network structure.

[22]Details can be found in the Appendix.

FIGURE 1.10: The Model Implied Evolution of the Phillips Multiplier in the U.S.: The Role of Sectoral Reallocation



Note: This figure shows the evolution of the model implied slope of the Phillips (black dashed line) together with a counterfactual slope that is derived from a multi-sector model with with no change in the value-added shares, ϑ_k, from 1963

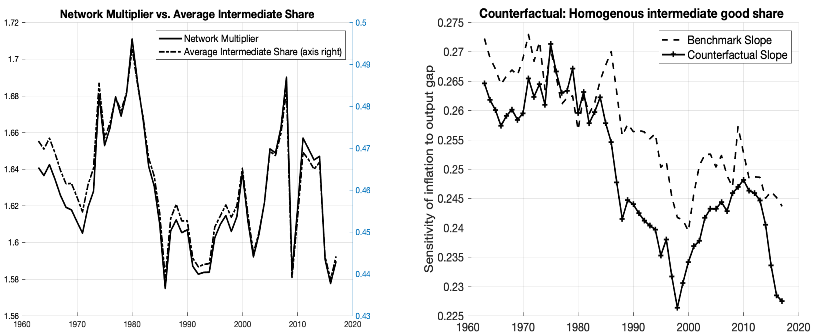
1.6.4 *Channel 2: The Role of the Network Multiplier.*

The second potential channel through which changes to the network structure can influence the network bias and, hence, the slope of the Phillips curve is via the network multiplier. According to equation (1.37) the most important determinant for the network multiplier is the output share weighted average intermediate input share, $\sum \delta_{k,t}\gamma_{k,t}$ because I abstract from heterogeneity in markups in the calibrated model.

Panel (a) of Figure 1.11 shows that the calibrated network multiplier is proportional to the average output share weighted intermediate good share. This also provides an external validity exercise since the model does a good job at matching the observed (and un-targeted) network multiplier from the data (Figure 1.5). Panel (a) of Figure 1.11 stresses the result from Section 1.4.2 that the network multiplier is relatively constant over time. The underlying reason for this is twofold and can be explained by observing the cross-section.

Figure 1.8 shows the sectoral change in the intermediate good use

Figure 1.11: The Model Implied Evolution of the Phillips Multiplier in the U.S.: The Role of the Network Multiplier



(a) Relationship Intermediate Share and Network Multiplier

(b) Counterfactual: Homogeneous Intermediate Good Shares

Note: Panel (a) of this Figure displays the average intermediate good share (black dashed line) and the network multiplier (black solid line). Panel (b) shows the evolution of the the model implied slope of the Phillips (black dashed line) together with a counterfactual slope that is derived from a multi-sector model with an homogeneous intermediate good share.

shares between 1963 and 2017. First, manufacturing firms (blue) tend to have larger intermediate good shares than service sectors (red). Therefore, the reallocation to service sectors should decrease the aggregate share of intermediate goods used in the economy. However, Panel B of Figure 1.8 documents that the intermediate good shares of service sectors (red) have increased from 1963 to 2017. The two effects counteract, and the overall intermediate share stays relatively constant.

To assess the role of the network multiplier on the slope, Panel (b) of Figure 1.10 reports the result of a counterfactual exercise. The Figure shows the implied slope of the Phillips curve (black dashed line) together with the implied slope derived from a model with homogeneous intermediate good shares. The evidence shows that heterogeneity in the degree of intermediate good shares cannot explain the flattening as the counterfactual slope decreases stronger than in the baseline case.[23]

1.7 CONCLUSION

A growing literature documents that the Phillips curve has flattened over time. I contribute to this discussion by providing evidence that changes to the network structure can be an important explanation. Using a multi-sector model with production networks, I show that input-output linkages in the production function of firms affect inflation dynamics and introduce a bias in the estimation of the slope of the Phillips curve. The size of the bias depends on the network structure and on the degree of strategic complementarities. When calibrated to the U.S. economy from 1963-2017, changes to the network structure of the economy are able to explain a significant part of the flattening of the Phillips curve. In this project, I abstracted from two dimensions of change in the network that are promising avenues for future research: (i) international linkages in production networks and (ii) the role of rising market power for the network multiplier. Accounting for production networks and its changes over time has important implications for inflation dynamics and contributes to better understand the changes to the Phillips curve relationship in the past decades.

[23] In the calibration exercise, variations in the network multiplier are solely caused by changes in the intermediate good share since we abstract from changes in markups. However, a number of recent studies have shown that market power may have risen in several sectors in the economy (DeLoecker et al., 2020). Heterogeneity in markups is an alternative way to explore variations in the network multiplier.

1.8 APPENDIX

EVIDENCE ON THE FLATTENING OF THE PHILLIPS CURVE

In this section, I present evidence on the evolution of the Phillips curve over the past 50 years. I will focus on three components: (i) the size of the change, (ii) the pattern of the change and (iii) the timing of the change.

Coefficient over Time

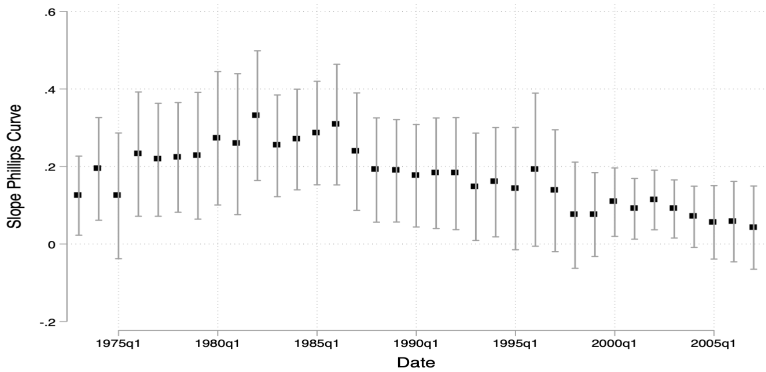
I continue by estimating this relationship for the U.S. economy between 1960 and 2007. These estimates are robust to different specifications, for instance concerning the measure of inflation, or the output gap. To characterize the strength of economic activity, I use estimates of the Congressional Budget Office for the potential level of GDP. Concerning inflation expectations, I follow Ball and Mazumder (2011) as well as Coibion and Gorodnichenko (2015), and assume as a simple baseline that inflation expectations are backward-looking. Specifically, I assume that inflation expectations are a four-quarter average of past inflation rates,

$$\mathrm{E}_t \pi_{t+1} = \frac{1}{4}(\pi_{t-1} + \pi_{t-2} + \pi_{t-3} + \pi_{t-4}).$$

where I use the inflation rate from the personal consumption expenditure survey (PCE), π_t.

First, I investigate how the sensitivity of inflationary dynamics to economic activity has changed over time. Was there a particular point in time when the slope broke down or was this rather a smooth process? Has the slope only flattened or was there a time when it was increasing? To answer these questions, I estimate the relationship (1.1) by OLS over rolling windows of 50 quarters.

FIGURE 1.12: Rolling Window Estimates of Phillips Curve Slope



Note: This figure displays the results of a rolling window estimates of the Phillips curve as in Equation (1.1). Window size is 50 quarters.

In Figure 1.12, I report the average relationship between the output gap and deviations of inflation from expectations for each window, κ, together with the one standard deviation confidence intervalls. Two results stand out. First, the slope of the Phillips curve has not always been flattening. Instead, we can basically observe three episodes since 1975. In the first part of the sample and up the middle of the 1980s, we can observe an increasing slope of the Phillips curve. In the second period which goes until the start of the 2000s, there is an apparent flattening of the slope. In the final phase, the relationship diverges on a low level, with periods in which the average relationship is significant and insignificant.

Second, there has not been a particular event that reduced the slope permanently. Instead, we can observe a protracted episode in which the slope has decreased since the mid-1980s. Therefore, the results of this section indicate that the flattening of the Phillips curve is a smooth process that started in the middle of the 1980s.

Timing of the Flattening

After identifying the shape of the flattening of the Phillips curve, I investigate the exact timing of the change. Therefore, I employ a Andrews (1993) test for parameter instability with unknown break date. I investigate the statistical evidence for a structural break in the relationship between inflation dynamics and the output gap by formally allowing for a break in the relationship at unknown τ as follows:

Table 1.3: Test for the Break Date in the Phillips Curve

Specification	Break Date	p-value
π^{PCE}, y_t	1982q3	0.0323
π^{CPI}, y_t	1983q2	0.0032
π^{PCE}, u_t	1981q1	0.0000
π^{PCE}, $ugap_t$	1982q3	0.1260
π^{CPI}, u_t	1995q3	0.0008
π^{CPI}, $ugap_t$	1983q2	0.0917
IV: π^{PCE}, y_t	1983q2	0.0502
IV: π^{PCE}, $ugap_t$	1982q1	0.0942
IV: π^{PCE}, ur_t	1982q1	0.0009

Note: This table displays the results of an Andrews (1993) test with 15% trimming and supremum LR-test.

$$\pi_t - \beta E_t \pi_{t+1} = c + \kappa_1 * I(t < \tau) * x_t + \kappa_2 * I(t \geq \tau) * x_t + v_t$$

where I are time dummy variables equal to one if the respective condition is satisfied and zero otherwise. The null hypothesis of the Andrews test is that the slope coefficients in both periods are equal.

Table 1.3, reports results for the Andrews test for different specifications for the inflation measure, forcing variable and estimation methodology. Consistent with results of the previous section, the test cannot reject the null that the slope is unchanged. The protracted flattening episode is represented by a series of potential break dates in 1982 or 1983.

Next, I will use the previous result on the break date to calculate the size of the change in the slope coefficient. In Figure 1.2, I present a scatter plot of quarterly output gaps for the United States against the deviations of inflation that quarter from expected, discounted inflation. Data from 1960Q1 until our estimated break date 1982Q3 is represented by circles, while data from 1982Q4 is plotted by diamonds. The lines represent the slope of the average relationship estimated by OLS over each sample period. The slope is positive, indicating that economic slack, i.e. economic activity that is lower than potential, is associated with inflationary pressures below expectations. The sensitivity has changed over time. We can observe a flattening in the sensitivity of inflation to economic activity and estimates of the slope suggest that those are in the magnitude of 40 to 80 percent.

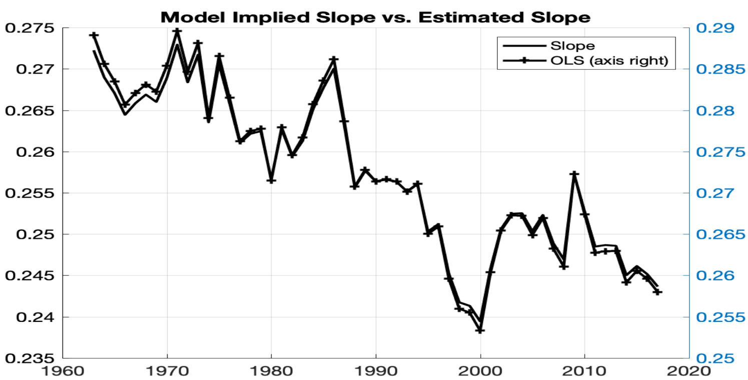
The statistical evidence of this section provided insights into the size, timing and shape of the flattening of the Phillips curve. The findings suggest that the flattening was a smooth process that started in 1982Q2 and saw a decrease in the slope of 40-80%.

The Bias from Inflation Expectations

In the monte carlo study in Section 1.5, there was a small difference between the OLS estimate of the slope, $\widehat{\kappa}$, and the sum of the standard slope and the biases. The difference is due to the unobserved inflation expectation. In this section, I quantify the size of the bias by comparing the model implied slope of the Phillips curve from simulated data from the multi-sector model to assuming that we can observe inflation expectations and estimate a modified projection on the model simulated data:

$$\widehat{\pi}_t^{Sim} - \beta\widehat{\pi}_{t+1}^{Sim} = \kappa * \widehat{y}_t + v_t.$$

Figure 1.13: Bias From Inflation Expectations in Estimating Phillips Curve Slope



Note: This figure compares the baseline estimate of the evolution of the model implied slope of the Phillips curve (black solid) with a counterfactuar exercise where we assume that the econometrician can observe inflation expectations without error (black solid with markers).

Figure 1.13 reports that the bias from not observing inflation expectations in the present exercise is small.

Relationship Between Network Statistics, Leontief Inverse and Domar Weights

In this section, I will show that the two network statistics introduced in this paper are related to other network statistics used in the literature: the Domar weights and the Leontief-inverse. First, the Domar weights are defined as the ratio of sectoral gross production to value-added output, which in the notation of this paper yields

$$\lambda_{k,t} = \frac{p_{k,t}Q_{k,t}}{GDP_t} = \frac{p_{k,t}Q_{k,t}}{Y_t}.$$

Second, the Leontief-inverse can be expressed as the infinite sum of the powers of the (adjusted) input-output matrix W (Carvalho and Tahbaz-Salehi, 2019)

$$L = [I_K - \gamma\frac{\overrightarrow{1}}{M_t}W]^{-1}.$$

Finally, the vector of Domar weights can be shown to be characterized by the Leontief-inverse or the two network statistics ((i) output share and (ii) network multiplier)

$$\Lambda = V_c' L = \delta\Phi^{NM}.$$

Alternative Examples of Output Structures

Example 3: Change in Output Share. In the next example, I want to show how the change of the I/O structure can change the output shares. Keeping the consumption shares, and intermediate good shares from Network 3 but increasing the markup from Sector 1 to 50%, i.e. $\frac{1}{\mathbb{M}} = [3/6; 5/6]$, I will look at the transition from Network 4 to Network 3, where the former is given by

$$W_4 = \begin{pmatrix} 1 & 0 \\ 0.5 & 0.5 \end{pmatrix}$$

In Network 4, Sector 2 is equally spending its expenditure on inputs from both sectors instead of solely using inputs from Sector 1 (Network 3). Therefore, Sector 1 will be relevant again but with a lower output share $\delta_4 = [1/3; 2/3]$. Transitioning to Network 3, the outdegree of Sector 1 increases and so does its output share to $\delta_3 = [1/2; 1/2]$. The increase in the centrality has another interesting implication in this example. Since, Sector 1 has a higher markup, the aggregate output share weighted markup increases. This decreases the network multiplier from 2.18 to 2.[24]

Example 4: Increasing the "Length" of the Production Chain. In the last example, I want to show how the change of the I/O structure can increase the network multiplier by increasing the length of the network. Keeping the consumption shares, and markups from Network 1, I will consider the following two networks

$$W_5 = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \text{ and } W_6 = \begin{pmatrix} 0.9 & 0.1 \\ 0.9 & 0.1 \end{pmatrix} \text{ and } \gamma_5 = \gamma_6 = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

[24] This is also smaller as the multiplier in the Network 3 in Example 2, of 2.67, because in Example 2 both sectors had lower markups.

While Network 5 is a symmetric network again as in example 1, Network 6 is a star network with Sector 1 being the central sector. Increasing the centrality of the sector that uses more intermediate goods can be seen as increasing the overall length of that network as discussed in Section 3.4.6. Making Sector 1 the “star” in the star network increases its centrality from 0.5 to 0.73 as measured by its output share. This increases the network multiplier from 2.08 to 2.32.

Log-lin system of equations

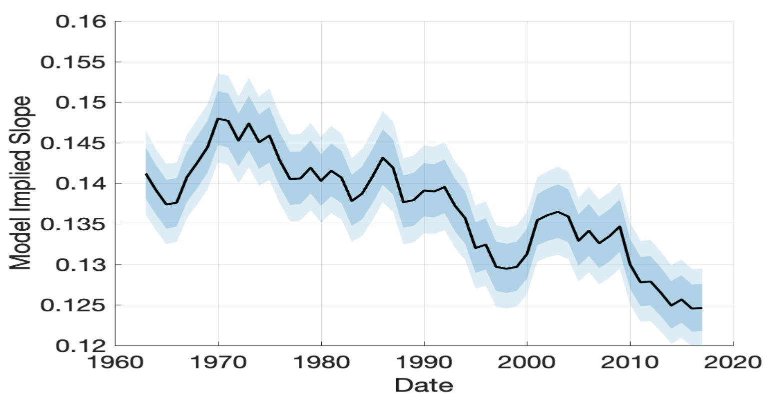
The log-linear system of equations is described by

$$
\begin{aligned}
\widehat{y}_t &= \mathrm{E}_t\widehat{y}_{t+1} - \frac{1}{\sigma}(\widehat{i}_t - \mathrm{E}_t\widehat{\pi}_{t+1}) \\
\widehat{i}_t &= \phi_y\widehat{y}_t + \phi_\pi\widehat{\pi}_t + \widehat{z}_t^m \\
\widehat{p}_{k,t} &= \widehat{\pi}_{k,t} - \widehat{\pi}_t + \widehat{p}_{k,t-1} \\
\widehat{\pi}_{k,t} &= \beta E_t\widehat{\pi}_{k,t+1} + \kappa_k\big(\frac{(1-\gamma_k)\varphi}{1+\gamma_k\varphi}(\widehat{p}_{k,t} + \widehat{q}_{k,t}) \\
&\quad + \frac{\sigma(1-\gamma_k)}{1+\gamma_k\varphi}\widehat{y}_t + \frac{(1+\varphi)}{1+\gamma_k\varphi}(\gamma_k\widehat{p}_t^k - \widehat{p}_{k,t})\big) \\
\widehat{p}_{k,t} + \widehat{q}_{k,t} &= \frac{\bar{Y}_k}{\bar{Q}_k}\widehat{y}_t + \sum_{r=1}^{K}\frac{\bar{X}_{r,k}}{\bar{Q}_k}(\widehat{mc}_{r,t} - \widehat{p}_{r,t} + \widehat{p}_{r,t} + \widehat{Q}_{r,t}) \\
\widehat{\pi}_t &= \sum_{k=1}^{K}\vartheta_k\widehat{\pi}_{k,t} \\
\widehat{p}_t^k &= \sum_{r=1}^{K}\omega_{k,r}\widehat{p}_{r,t}
\end{aligned}
$$

Accounting for the Fall in the Labor Share

In this section, I recalibrate the intermediate share to account for the fall in labor share. Because I assume only two inputs in production but the BEA input-output tables consider three inputs, there is an ambiguity in the calibration. Instead of calibrating the share of intermediate goods used, I target the labor share. This results in a fall of the labor share and an increase in intermediate shares in the economy over time. However, while the original calibration is consistent with the observed network multiplier, this alternative calibration results in an increase in the network multiplier over time. Figure 1.14 shows the result of this alternative calibration. The slope is falling by about 18%. The fall is stronger because of an increasing network multiplier, and, thus, bias from the network component in the model implied slope.

Figure 1.14: The Model Implied Slope of the Phillips Curve With a Falling Labor Share



Note: This figure shows the model implied slope of the Phillips curve under an alternative calibration that targets the labor share instead of the intermediate good share.

More Evidence on Changes in the Production Network Over Time

Finally, in this section, I will provide additional evidence on how the distribution of industries has changed in terms of value-added shares, output shares and intermediate good shares. In general, we can observe again structural changes in the network structure (process of service deepening) additionally to structural transformation.

Table 1.4: Comparison of Most and Least Central Sectors in 1963 vs. 2017

Top 10 central industries in 1963		Top 10 central industries in 2017	
23	Construction	531	Real estate
311FT	Food, beverage, tobacco products	5412OP	Manage Prof scientific technical
44RT	Retail trade	42	Wholesale trade
531	Real estate	81	Other services, except government
81	Other services, except government	44RT	Retail trade
5412OP	Manage Prof scientific technical	622	Hospitals
42	Wholesale trade	23	Construction
3361MV	Motor vehicles, and parts	524	Insurance and related
111CA	Farms	621	Ambulatory health care services
325	Chemical products	311FT	Food, beverage, tobacco products
Bottom 10 central industries in 1963		**Bottom 10 central industries in 2017**	
514	Information services	486	Pipeline transportation
486	Pipeline transportation	315AL	Apparel and leather products
493	Warehousing and storage	483	Water transportation
562	Waste management services	323	Printing and support activities
213	Support activities for mining	337	Furniture and related products
483	Water transportation	313TT	Textile mills
523	Securities	482	Rail transportation
721	Accommodation	212	Mining, except oil and gas
323	Printing and support activities	113FF	Forestry, fishing, and related
481	Air transportation	562	Waste management

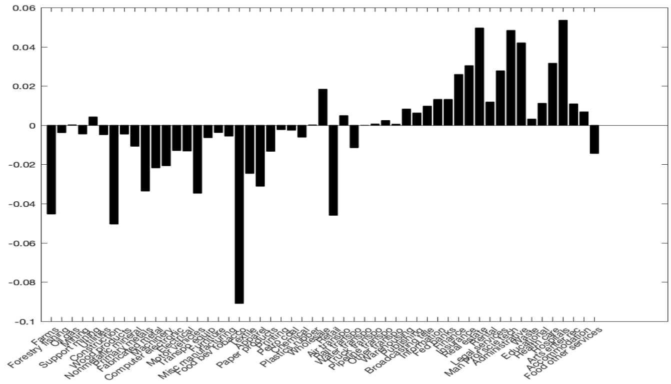
Note: This table compares sectors' centrality in 1963 vis-a-vis 2017. Bureau of Economic Analysis and author's calculations.

Table 1.5: Characteristics of Most Central Sectors

	Industry	θ_k	γ_k^{1963}	γ_k^{2017}	ΔW_c	$W_k^{',1963}$	$W_k^{',2017}$
Top 10 rank improvements from 1963 to 2017							
622	Hospitals	0.79	0.24	0.45	0.05	0	0.01
531	Real estate	0.50	0.13	0.31	0.01	1.61	3.26
5412OP	Manage Prof scientific technical	0.78	0.37	0.37	0.00	3.46	6.17
621	Ambulatory health care services	0.83	0.39	0.34	0.03	0.01	0.07
561	Administrative and support services	0.65	0.22	0.40	0.00	0.64	2.96
524	Insurance carriers and related activities	0.82	0.49	0.57	0.01	1.53	1.58
5411	Legal services	0.87	0.45	0.27	0.02	1.05	1.46
523	Securities	0.55	0.53	0.47	0.01	0.22	0.99
42	Wholesale trade	0.45	0.16	0.38	0.01	2.42	3.70
521CI	Federal Reserve banks	0.08	0.37	0.31	0.01	1.75	1.53
Top 10 rank decline from 1963 to 2017							
	Industry	θ_k	γ_k^{1963}	γ_k^{2017}	ΔW_c	$W_k^{',1963}$	$W_k^{',2017}$
311FT	Food. beverage. tobacco products	0.61	0.71	0.72	-0.04	1.63	0.85
23	Construction	0.42	0.59	0.49	-0.05	0.87	0.56
44RT	Retail trade	0.56	0.41	0.39	-0.04	0.69	0.46
111CA	Farms	0.02	0.52	0.65	-0.01	1.15	0.68
3361MV	Motor vehicles and parts	0.69	0.64	0.78	-0.02	1.08	0.81
315AL	Apparel	0.81	0.69	0.64	-0.02	0.5	0.35
331	Primary metals	0.20	0.61	0.72	0.00	2.64	1.64
333	Machinery	0.78	0.46	0.60	-0.01	0.96	0.85
313TT	Textile mills	0.75	0.70	0.66	0.00	1.32	0.59
332	Fabricated metal products	0.73	0.54	0.58	0.00	1.76	1.62

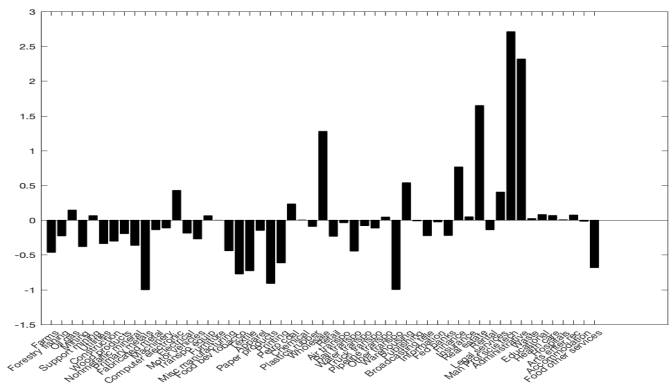
Note: This table compares the change in sectors' centrality from 1963 to 2017 and reports some key characteristics such as frequency of price change θ, intermediate share γ, change in value added share ΔW_c, and outdegree W'. Bureau of Economic Analysis and author's calculations.

Figure 1.15: Change in Sectoral Output Shares, $\delta_{k,t}$, from 1963 to 2017



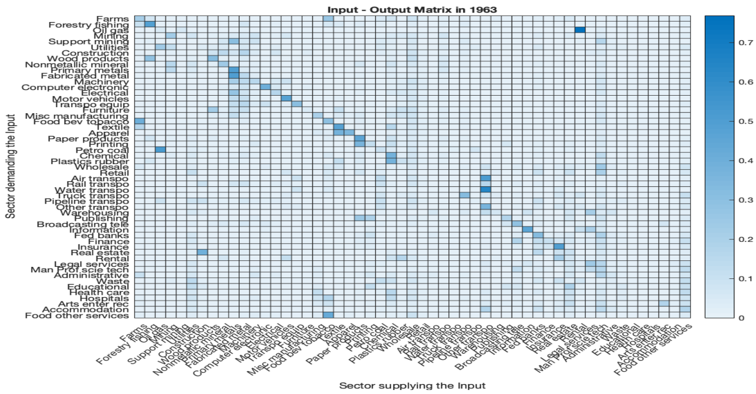
Note: This figure shows the change in sectoral output shares, $\delta_{k,t}$, a measure of the network centrality of a sector. Services became the most central sector in the economy. Sources: Bureau of Economic Analysis and author's own calculations.

Figure 1.16: Change in Sectoral Outdegrees from 1963 to 2017



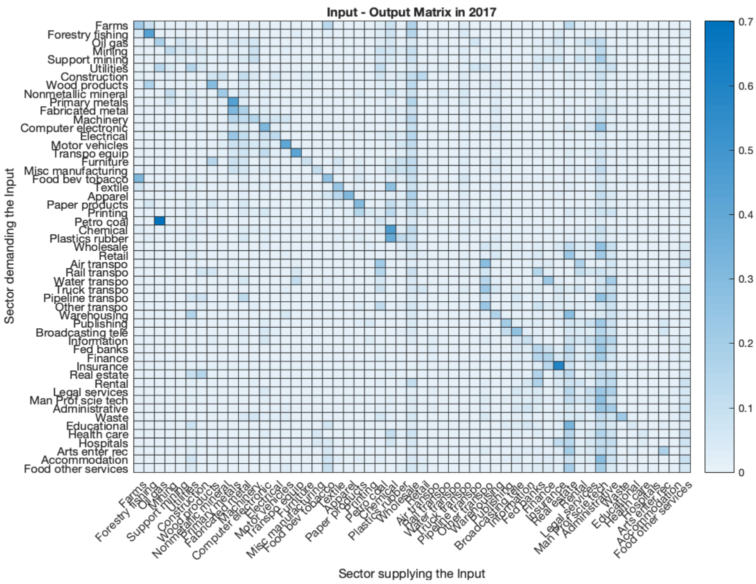
Note: This figure displays the change in sectoral outdegrees, which are another measure of network centrality and equal to W'. Services became the most central sector in the economy also according to the outdegree. Sources: Bureau of Economic Analysis and author's own calculations.

Figure 1.17: Input-Output Matrix in 1963



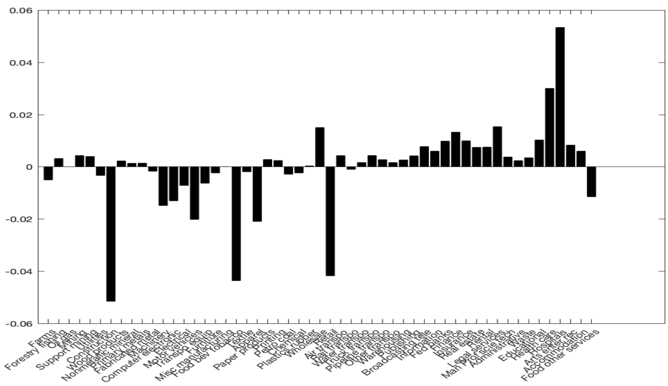
Note: This figure displays the Input-output matrix as created by the guidelines of BEA handbook in 1963. It shows how much inputs each sector is demanding from other sectors (vertical axis) and how much it is providing (horizontal axis). The shades of blue represent the proportion of use, whereby darker shades represent more intense use.

FIGURE 1.18: Input-Output Matrix in 2017



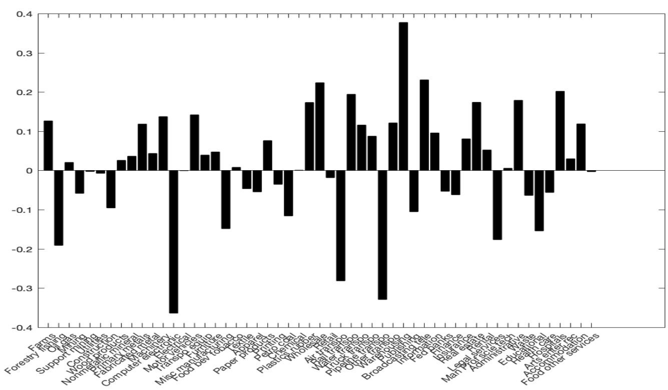
Note: This figure displays the Input-output matrix as created by the guidelines of BEA handbook in 2017. It shows how much inputs each sector is demanding from other sectors (vertical axis) and how much it is providing (horizontal axis). The shades of blue represent the proportion of use, whereby darker shades represent more intense use.

Figure 1.19: Sectoral Distribution of Change in Value-Added Share from 1963 to 2017



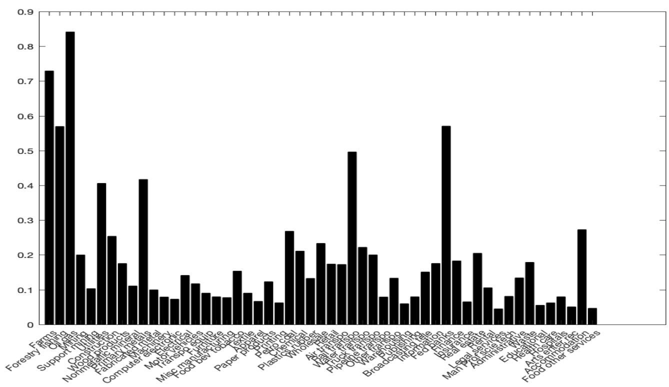
Note: This figure illustrates the change in sectoral value-added shares as measured by final demand from the BEA Use table from 1963 to 2017. While manufacturing sectors see a decline in the importance, services increase strongly in terms of GDP share.

Figure 1.20: Change in Sectoral Intermediate Good Use Share



Note: This figure shows the change in sectoral intermediate good use in industries' production from the BEA Use table from 1963 to 2017.

FIGURE 1.21: Sectoral Distribution of Frequency of Price Adjustment



Note: This figure plots the distibution of frequency of price change for each sector. It is measured as the average proportion of goods within each sector from the BEA survey of PPI firms that changes prices each month. I thank Michael Weber for providing the data. Generally manufacturing sectors tend to have a higher frequency of price change, i.e. they are less rigid than services.

2

DISPERSED MARKET POWER, PHILLIPS MULTIPLIER, AND THE OPTIMAL INFLATION TARGET

(joint with Donghai Zhang)

Several indicators suggest that competition may be decreasing in many economic sectors, including the decades-long decline in new business formation and increases in industry-specific measures of concentration.

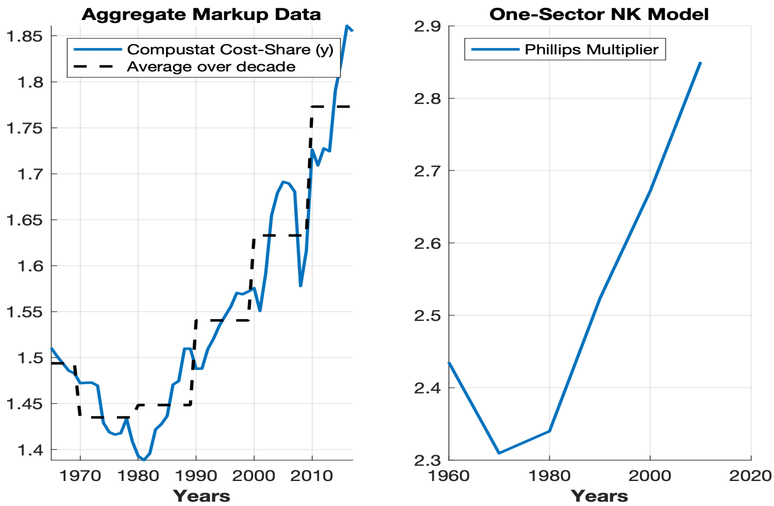
— Council of Economic Advisors, 2016

2.1 INTRODUCTION

Recent contributions in empirical macroeconomics have highlighted that the average markup/market power for firms in the U.S. has increased over the past decades (DeLoecker et al., 2020). Market power has an interesting interaction with nominal rigidities if firms have non-constant returns to scale production function. According to a basic one-sector New Keynesian model with decreasing returns to scale, this increase in markups implies an increasing Phillips Multiplier and a decreasing monetary non-neutrality (see, e.g., Coibion and Gorodnichenko, 2015), see Figure 2.1. The intuition is the following. Facing a reduction in marginal cost, a firm resets its price downward. Due to staggered prices, such a price cut then generates excess demand in the future. With decreasing returns to scale, marginal costs would rise in the future. As a result, the firm cuts its price less than it would otherwise do in the absence of this feedback effect. The degree of competition in the market amplifies this effect, altering the effects of real shocks to nominal prices. The previous analysis ignores the possibility that (i) the distribution of the average markups are dispersed and that (ii) the entire distribution might be evolving over time. How has the entire distribution of markups evolved over time in the data? What are the implications for the conduct of monetary policy?

In this paper, we first complement the recent empirical literature by showing that the distribution of steady-state markups spread out over time. This is driven by both the increase in the top quantiles and a decrease in the bottom quantiles of the distribution. We then study the implications of those findings based on a New Keynesian model with heterogeneous sectors. Particularly, we examine the implications of dispersed markups and the evolution of the distribution of markups over time for (i) monetary non-neutrality, (ii) the Phillips Multiplier, and (iii) the optimal inflation index (OII) stabilization policy. Understanding the degree of monetary non-neutrality, the size of the Phillips Multiplier,

FIGURE 2.1: Evolution of U.S. Markups and its Implications



Note: Authors' own calculation. The solid blue line in the left panel plots the aggregate markup computed using the cost-share approach using data from Compustat that covers publicly listed firms in the U.S. The dashed black line reports the average over the respective decade, which we interpret as the markup's steady-state value. The right panel reports the implied Phillips Multiplier for each decade based on the simple New Keynesian model outlined in Gali (2015).

and the composite of the OII are essential for the conduct of monetary policy. In fact, these three statistics are the foundations of the Federal Reserve's (Fed) dual mandate: Foster economic conditions that achieve both stable prices and maximum sustainable employment. The first statistic, monetary non-neutrality, measures the central bank's ability to stimulate the economy to achieve maximum sustainable employment. The second statistic, the Phillips Multiplier (Barnichon and Mesters, 2020b) – defined as the ratio of the cumulative response of inflation to the cumulative response of real GDP after an exogenous monetary intervention – measures the trade-off between the stabilization of prices and the stimulation of employment. Third, the OII informs policymakers which inflation index they should target to achieve price stability and whether their current practice is close to the optimal. In the U.S., the Fed monitors the headline and the core of the personal consumption expenditures (PCE) inflation, which does not necessarily coincide with the OII. Our contribution is in showing how dispersed steady-state markups affect those three statistics both on average and over time are relevant for guiding policy discussions.

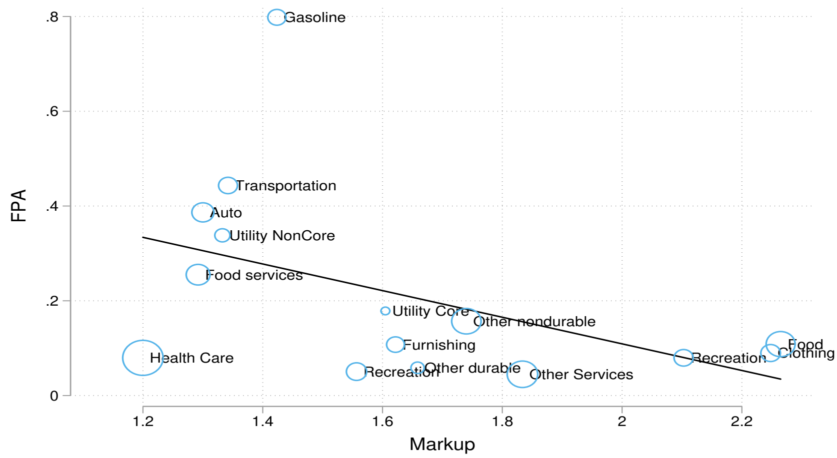
We derive the following results. First, dispersed steady-state markups lead to stronger money non-neutrality, and a smaller Phillips Multiplier in the presence of decreasing returns to scale. This is because monetary non-neutrality is concave in the steady-state level of markups. This result suggests that the central bank's ability to stimulate employment might be higher than previously thought, based on a model with homogenous market power or constant returns to scale.

Second, we investigate how changes in steady-state markup distribution and sector sizes affect the size of the Phillips Multiplier based on calibrations of a seventeen-sectors model that consists of constitutes of the PCE index in the United States. We find that changes in steady-state markup distribution have a minimal impact on the size of the Phillips Multiplier independent of the returns to scale. This result is driven by the spreading out of the markup distribution: Effects that arise through the changes in the right tail cancel out with the effects emerging from the movements in the left tail. However, our calibrated model predicts a 20% reduction in the Phillips Multiplier due to the reallocation of resources to stickier price sectors.

We then study the policy implications. Specifically, we study the OII stabilization policy: the optimally weighted inflation index that the central bank should target to minimize the social welfare loss. One important mechanism that academics and policymakers focus on is the relative price stickiness channel. We label this as the *stickiness channel*, see e.g., Aoki (2001), Benigno (2004) and Mankiw and Reis (2003). In this paper, we address the heterogeneity in markups (the *competition* *channel*) in the design of the OII.

Empirically, we document that markups are negatively correlated with nominal rigidities across sectors (see Figure 2.2), which is consistent with costly price adjustment models developed by Barro (1972), Sheshinski and Weiss (1977) and Golosov and Lucas (2007). Therefore, analyzing the stickiness channel without considering the origin of the relative frequency of price adjustment might be misleading.

FIGURE 2.2: Frequencies of Price Adjustment vs. Markups



Note: Authors' own calculation. This figure plots the frequencies of price adjustment against steady-state markups in seventeen sectors that are constitutes of the PCE index in the U.S. The size of a circle measures the size of the underlying industry. The black line is the fitted linear relationship according to the OLS.

We show that a more competitive (lower market power) sector is associated with a higher weight in the OII. In the extreme case, when a market is infinitely close to a perfect competition market (flat demand curve), the optimal inflation index is the one that only consists of inflation in that sector. The intuition is as follows. In a more competitive market, firms face a flatter demand curve. Consequently, a given change in prices leads to a more significant movement in quantity. In the presence of price stickiness, this results in a more significant dispersion in output, which is welfare detrimental due to consumers' love of variety. In sum, inflation in a more competitive sector creates a bigger distortion. Therefore, stabilizing inflation in that sector is relatively more important, hence the higher weight. To illustrate the interaction between the competition channel and the stickiness channel, we calibrate a two-sector model with heterogeneous degrees of nominal rigidities and market power to the manufacturing and service sectors in the data. Interestingly, the competition channel offsets the stickiness channel. As a result, the PCE (weighted by the size of the market) stabilization performs similarly as compared to the stabilization of an inflation index that is merely based on the relative price stickiness. We label the latter as the stickiness-based price index (SPI). This finding challenges the conventional wisdom that the central bank should always attach a higher weight to a sector with a higher degree of nominal rigidity.

We compute the OII for the seventeen-sector model calibrated to the U.S. data over time. In the 1960s, the competition channel played a minimal role: the welfare loss associated with the stabilization of the SPI is almost identical to the case of the stabilization of the OII. However, changes in the distribution of market powers that have occurred in the data affected this result. In the twenty-first century, ignoring the heterogeneity in market competition results in a welfare loss that is 6.1%, measured in terms of welfare loss under a PCE stabilization policy, higher than the outcome under the OII stabilization policy.

Lastly, we conduct a positive analysis by plotting the OII and compare it with the headline and the core PCE. A simple visual inspection suggests that during the Great Moderation periods, the OII was consistently higher than the two PCE measures that the Fed relies on in their policy analysis. This demonstrates that the OII stabilization cannot be achieved by monitoring a weighted average of the headline and the core PCE. During the periods following the Great Recession of 2008, similar to other inflation measures, the OII is below the 2% target.

Literature review This paper is related to studies on multi-sector New Keynesian models. Those studies share the insight that heterogeneous price rigidity increases the effects and persistence of demand shocks, e.g., Carvalho (2006), Nakamura and Steinsson (2010), or Carvalho and Schwartzman (2015). In recent studies, the focus has shifted to the interaction of nominal rigidities with other sources of heterogeneities. Pasten et al. (2019) or Chapter 1 show that production networks can magnify the importance of price rigidities through its effects on marginal costs. In contemporary work, Reinelt and Meier (2020) show that firms with more rigid prices optimally set higher markups due to the precautionary price-setting motive. We study the implications of steady-state markup dispersions for the conduct of monetary policy, and we highlight the importance of monitoring the entire distribution.

Previous literature on the optimal inflation index is abundant, but most conclusions are drawn based on frameworks that introduce nominal rigidity into different markets, in the spirit of Aoki (2001), Benigno (2004) and Mankiw and Reis (2003). Erceg et al. (2000) show that in the presence of nominal wage rigidity, the optimal monetary policy index includes wage inflation. Huang and Liu (2005) demonstrate that with price stickiness in intermediate sectors, it is optimal for the central bank to respond to both PCE inflation and PPI inflation. By introducing nominal rigidity to the investment goods sector, Basu and Leo (2016) conclude that the optimal policy reacts to inflations in both consumption goods and investment goods. Anand et al. (2015) consider the optimal inflation targeting policy for developing countries. They show that with a significant fraction of hand-to-mouth workers in the food sector, stabilizing headline PCE is welfare improving compared to maintaining core PCE. Eusepi et al. (2011) derive an optimal inflation index considering heterogeneity in nominal rigidity and the labor share and find that optimal weights mostly depend on price stickiness. We show that a sector with a more rigid price is not necessarily associated with a higher weight in the OII due to the competition channel and its empirical correlation with the stickiness channel.[1]

2.2 EMPIRICAL EVIDENCE

This section documents new empirical observations on the dispersion of markups and the empirical relationship between nominal rigidity and market power across sectors. In detail, we show additional facts regarding to the increase in average market power over time: (i) the dispersion in markups increases over time, (ii) this is driven not only by increases in the markups of high markup firms but also by decreases for low markup firms. Additionally, we show that (iii) firms in more competitive sectors change prices more often.

2.2.1 *Data*

Before we turn to the evidence, we first outline the data we use for the empirical analysis and in order to calibrate our theoretical model. Specifically, we combine and match data from three different data sources.

Firm-Level Markups. We use quarterly firm-level balance-sheet from 1967 - 2017 of publicly traded firms in Compustat to calculate firm-level markups. The data covers sales, employment, capital, and input factors of firms (cost of goods sold) over a long sample for a wide range of sectors covering manufacturing and service sector firms. We estimate firm-level markups following the single-input approach of Hall

[1]More broadly, this paper is related to the literature that studies the optimal monetary policy with a dynamic price elasticity originating from firm entry and exit. See, for example, Bilbiie et al. (2008), Bilbiie et al. (2014), Bergin and Corsetti (2008), Cooke (2016), Etro and Rossi (2015), Faia (2012) and Lewis (2013). In contrast to those studies, this paper focuses on the heterogeneity in the *steady-state* price elasticity across sectors. In another closely related paper, Andrés et al. (2008) rely on cross-country heterogeneity in competition to explain inflation differentials in the EMU.

(1986) and Hall (1988) and DeLoecker and Warzynski (2012). According to this approach, the markup $\mu_{i,t}$ of a firm i at time t can be computed from one flexible input, X_i, as the ratio of the output elasticity of the input, ε_{Q,X_i}, to the revenue share of that input, s_{R,X_i}

$$\mu_{i,t} = \frac{\varepsilon_{Q,X_i,t}}{s_{R,X_i,t}}. \tag{2.1}$$

Compustat reports a composite input called Cost of Goods Sold (COGS), which consists of intermediate and labor input and that will be used as the (partially) flexible input, X_i. DeLoecker et al. (2020) use a variant of the technique introduced by Olley and Pakes (1996) and described in DeLoecker and Warzynski (2012) to estimate a Cobb-Douglas function and obtain a time-independent estimate of output elasticity at the sector level. The markups are then derived by dividing the former by the share of COGS to revenue. We split our analysis into two parts given the critique on estimating markups using the production approach using revenue data, e.g., Bond et al. (2020) or Basu (2019). First, in the main part of the paper, we follow the cost share approach and focus on revenue shares to learn about the variation in markups across firms and over time. We calculate firm-specific revenue shares as the ratio of costs of goods sold to sales. Bond et al. (2020) outline that this approach can be used to study this variation under minimal restrictions without estimating an output elasticity. We check the robustness of these results to calculating markups as in DeLoecker et al. (2020) and report those results in the Appendix 2.7.

When transforming the data, we drop all firms in the sectors government or FIRE. We consider only observations that are positive and linear interpolate observations that are missing for one period. Additionally, we perform outlier adjustments by trimming at 1% (5%) of calculated markups.

One concern with Compustat is that it covers only publicly traded firms and thus is not representative of the distribution of the universe of firms. We account for a representativeness bias by using each sector's weights in the Compustat data from the PCE expenditure shares to account for sectoral composition (while we still calculate markups from publicly traded firms).

Frequency of Price Adjustment. We use sector-level frequencies of price adjustment, FPA, from producer price data (PPI), averaged over the period 2005 - 2011 from Pasten et al. (2019). The PPI measures selling prices of goods from the perspective of producers and covers goods-producing industries and services. The original confidential micro price data underlies the PPI and is collected by the BLS, covers about 25,000 establishments for approximately 100,000 individual items every month. The data we use is the median monthly frequency at the 6-digit NAICS level.

The aggregate price adjustment frequency, FPA, of matched industries is 0.64, close to the reported price frequencies in Nakamura and Steinsson (2008) or Bils and Klenow (2004). We can match 88 percent of firms with an FPA. Wherever we use it, we define implied price duration following Nakamura and Steinsson (2008) as -1/ln(1-FPA). In contrast to Nakamura and Steinsson (2008) or Bils and Klenow (2004), the data source is not consumer prices but producer prices. This aims to account for the fact that, first, markups are set at the producer level, and, second, the Compustat data is defined on NAICS levels.[2]

Personal Consumption Expenditures and Sectoral Prices. The use of personal consumption expenditure (PCE) data has two advantages. First, the main inflation target for monetary policy in the United States is the PCE deflator. Thus, we will use the PCE deflator as the reference

[2]Eusepi et al. (2011) aggregate FPA of entry-level items (ELIs) in the non-shelter component of the consumer price index (PCE) from Nakamura and Steinsson (2008) into PCE sectors. As we show when we discuss our seventeen sectors economy, when we aggregate the PPI based FPAs into PCE sectors, we arrive at mostly similar frequencies.

point for our optimal policy analysis. Second, we use PCE bridge tables to match NAICS sectors with the sectors that compromise the PCE deflator. For the baseline calibration, we choose 17 sectors. This choice is driven by the 15 major types of PCE products plus a division of utilities into a core and non-core component. This allows us to compare the resulting index to the core price PCE index. Moreover, we separate food services from accommodations since they have a relatively large share and are characteristically very different. We measure the size of sectors by their average expenditure share over each decade from NIPA Table 2.3.5.U. that reports personal consumption expenditures by Major Type of Product and by Major Function. We use PCE Bridge tables from the underlying detail estimates of the Industry Economic Accounts. These are annual tables that outline the commodity composition of the PCE categories from the National Income and Product Accounts (NIPAs) from 1997-2019.[3] In detail, they specify for each PCE category the commodities it is composed of together with the purchasers' value, which we will use as weights. We will use this information to match each PCE sector into the different NAICS sectors for which we have generated estimates on markups and price rigidity.[4] Finally, we aggregate the matched data into three different levels: a one-sector economy, a two-sector economy – in which we distinguish between goods and services – and a 17-sector case composing of the major types of products of the PCE. To study the optimal price index, we use sectoral price indices data from the underlying detail table 2.4.4.U. for personal consumption expenditures by type of product.

[3]For the years before 1997, we do not have information on the weights of commodity composition. We deal with this by considering the average weights between 1997-2019 for all years.

[4]Results of this exercise can be seen in Table 2.2

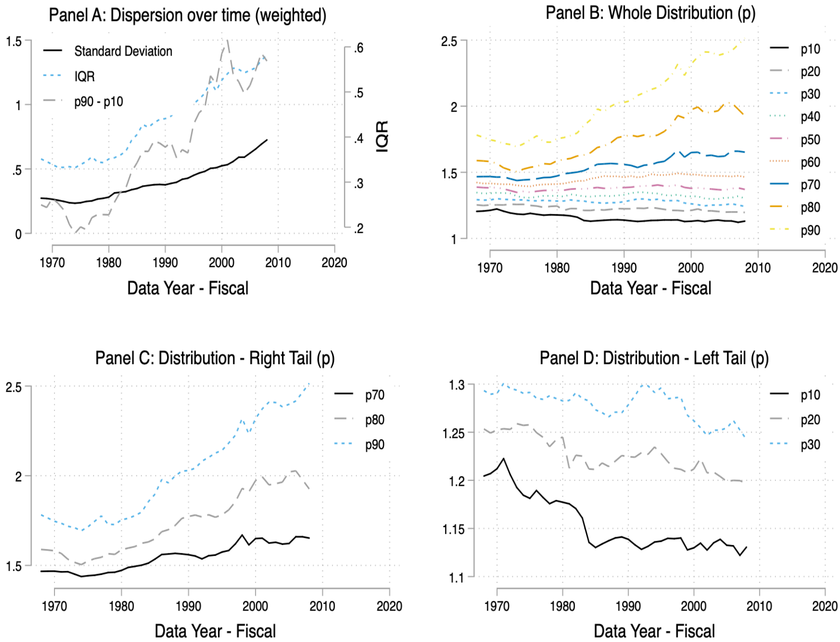
2.2.2 *Markup Dispersion and Correlation*

We derive new empirical results on the evolution of markups over time and their relationship to firms' price-setting behavior. Our focus is twofold. First, we want to document the evolution of the distribution of markups over time. Since we later uncover the consequences of different parts of the markup distribution on monetary policy, we can use those results to draw conclusions about alterations to the transmission mechanism induced by these changes. Second, we will use the cross-sectional variation to calibrate a 17-sector version of our theoretical model.

Observation 1: Markup Dispersion Increases Over Time. We estimate yearly markups at the firm-level from 1967-2017. We take 10-year moving averages of markups and obtain a distribution of smooth markups for each year. Panel A in Figure 2.1 plots the resulting average markup over time. The increase in markups compares to other recent estimates in a literature that documents increasing market power in economies. DeLoecker et al. (2020) compares the distribution of markups in 1980 and 2016 and argues that a thicker right tail – more mass of firms with high markups – leads to the higher estimate of the average markup over the sample time.

Based on the distribution of smooth markups over time, we then calculate the dispersion of individual markups in each year via three measures: (i) the standard deviation (ii) the interquartile range (IQR), and (iii) the range between the 90th and 10th percentiles. Panel A of Figure 2.3 and Figure 2.4 show that the resulting dispersion is increasing over time, independent of the considered measure. For all three measures, this process began in the 1980s and quantitatively led, for instance, to more than a doubling of the standard deviation. As for the two measures of range, we can observe an additional acceleration at the end of the 1990s. Moreover, the interquartile range (right y-axis) increases less,

Figure 2.3: Steady-State Markups by Quantiles and their Dispersion

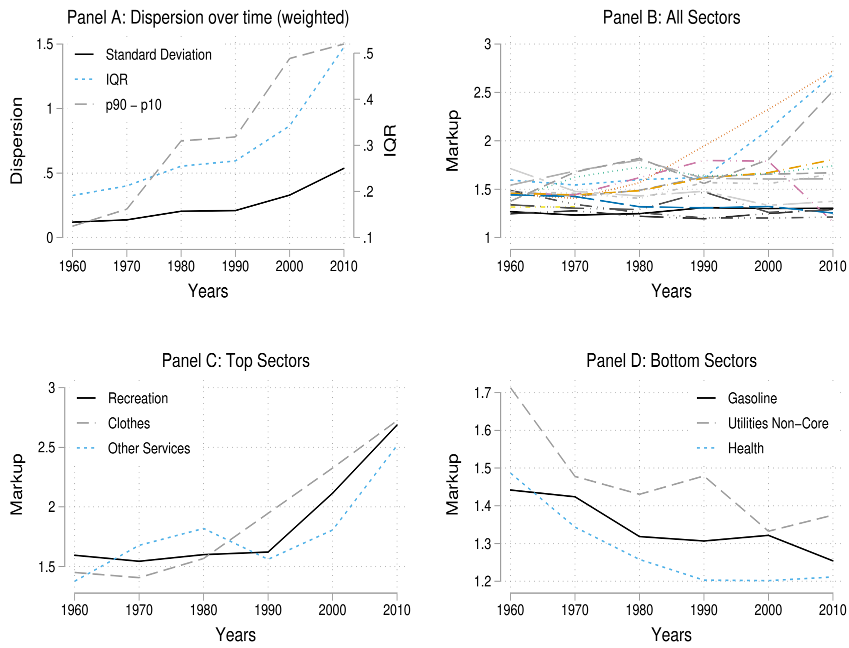


Note: Authors' own calculation. This figure displays the dispersion and the different quantiles of firm-level steady-state markups (measured as cost-share) in Compustat from 1967-2017.

reflecting the stronger increasing dispersion between very high markup firms and very low markup firms. Next, we want to investigate where the increase in dispersion is coming from.

Observation 2: Gap Between Left and Right Tail Widens by Both Sides. Where does the dispersion come from? Conceptually, it could arise because markups of high markup firms increase, due to falling markups at the left tail of the markup distribution, or both. DeLoecker et al. (2020) report that the mean markup increase is due to composition effects. Firms with larger markups increase in size. This

Figure 2.4: Steady-State Markups by Sectors and their Dispersion



Note: Authors' own calculation. This figure displays the dispersion and the steady-state markups (measured as cost-share) in Compustat from 1967-2017 for seventeen sectors that are constituents of the PCE.

leads to increases in the highest percentiles of the distribution.

Panels B and C of Figure 2.3 and Figure 2.4 confirm these results. For the highest percentiles (80 and 90), markups increase by more than 50 percent since the 1980s. In contrast, we observe decreases at the bottom of the markup distribution in smaller proportions. They are not sufficient to counteract the overall increase in average markups. However, as we will argue, they will become important later due to non-linearities in theoretical models even if their relative size appears to be irrelevant compared to the impressive increases at the top. We redo the same exercise for the seventeen sectors that constitute the PCE index. Figure 2.4 shows that we can find the same results on the dispersion but also on the gap between the left and the right tail in the less aggregated data.

In summary, we observe large increases in markups' dispersion across firms in the U.S. since the 1980s. We find that these are not only because of increases at the right tail of the distribution but also due to decreasing markups for low markup firms.

Observation 3: Firms With Higher Markups Change Prices Less Often. We also compare markups and nominal rigidity. In the data, we observe that firms can keep prices unchanged for an extended period of time. The degree of price rigidity is a leading explanation for the large effects of demand shocks (e.g., monetary policy) on output and is a central ingredient in New Keynesian macroeconomics. To test the relationship, we observe matched markup-frequency pairs at the 3-digit industry level from our constructed dataset. Since frequencies have been calculated over the 2005-2011 period, we calculate sectoral markups over the same timespan. In the first piece of evidence, we aggregate these 3-digit sectors into seventeen sectors that composite the PCE price index. Figure 2.2 illustrates a clear negative correlation between markups and frequency of price adjustment. The interpretation is that firms in sectors that are more competitive change prices more frequently. This is confirmed by a negative slope coefficient of an OLS regression on these seventeen sectors. We further verify this result by running OLS regressions with controls on a panel of 3-digit sectors. We find a negative and significant relationship between markups and the frequency of price adjustment in all versions. Details of this exercise are in the Appendix 2.7.

2.3 THE ECONOMIC MECHANISM

In the previous section, we documented that the steady-state markups are dispersed and evolving over time. We highlight both the positive trend in the right tail and the negative trend in the markup distribution's left tail. We will now assess the implications of those facts for the conduct of monetary policy. Before moving to the full model, we illustrate the key mechanism based on a basic NK model borrowed from Gali (2015) Chapter 3. For details about the setup and meanings of parameters, we refer readers to the original textbook. The following equations characterize the equilibrium of the economy:

$$\widetilde{y}_t = \mathbb{E}_t \widetilde{y}_{t+1} - \frac{1}{\sigma}[i_t - \mathbb{E}\pi_{t+1} - \rho)], \tag{2.2}$$

$$\pi_t = \beta \mathbb{E}\pi_{t+1} + \kappa \widetilde{y}_t, \tag{2.3}$$

$$i_t = \rho + \phi_\pi \pi_t + \phi_y \widetilde{y}_t + v_t, \phi_\pi > 1, \tag{2.4}$$

where $\widetilde{y}_t, \pi_t, i_t$ denote the output gap, inflation, and the nominal interest rate, respectively. Monetary shocks v_t, which follow a AR(1) process $v_t = \rho_v v_{t-1} + \varepsilon_t^v$, are the only shocks that hit the economy. The slope of the Phillips Curve κ is a composite of the deep parameters in the model:

$$\kappa \equiv \frac{(1-\beta\theta)(1-\theta)}{\theta} \frac{1-\alpha}{1-\alpha+\alpha\varepsilon} \left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right), \tag{2.5}$$

where the key parameter ε that we are interested in is the elasticity of substitution across goods. It is worth emphasizing that ε is the measure of the degree of competition in the economy, i.e., it is negatively related to the degree of market power. Specifically, the steady-state markup is $\frac{\varepsilon}{\varepsilon-1}$. The parameter α determines the returns to scale of the production.

The basic model can be solved analytically to obtain:

$$y_t = \widetilde{y}_t = -\Omega v_t, \tag{2.6}$$

where $\Omega \equiv \frac{1-\beta\rho_v}{(1-\beta\rho_v)[\sigma(1-\rho_v)+\phi_y]+\kappa(\phi_\pi-\rho)}$ measures the size of the effect of a monetary policy shock on real GDP, which we denote as the degree of monetary non-neutrality. From this expression, we can derive the following proposition.

Proposition 2.3.1 *With decreasing returns to scale, the slope of the Phillips Curve κ is a decreasing convex function of the elasticity of substitution across goods ε and the degree of monetary non-neutrality (Ω) is an increasing concave function of ε.*

Jensen's inequality implies that with decreasing returns to scale, the average of the money non-neutralities in different economies with heterogeneous market powers is higher than the money non-neutrality of a representative economy featuring the average market power. Similarly, the average of the slopes of the Phillips Curve in different economies with heterogeneous market powers is smaller than an economy with the average market power.[5]

The intuition for the slope of the Phillips Curve changing in ε is the following. In the presence of nominal rigidity, a firm's optimal price does not depend on the current marginal cost, but also the future ones:

$$p_t^* = \mu + (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k E_t\{\psi_{t+k|t}\}, \tag{2.7}$$

where ψ_t denote the log marginal cost, and $\psi_{t+k|t}$ denotes the marginal cost in period $t+k$ for a firm that last reset its price in period t. Moreover, the following relationship holds:

$$\psi_{t+k|t} = \psi_{t+k} - \frac{\alpha\varepsilon}{1-\alpha}(p_t^* - p_{t+k}). \tag{2.8}$$

This equation states that the marginal cost in period $t+k$ for a firm that last reset its price in period t is decreasing in p_t^*, as long as the marginal

[5] Imbs et al. (2011) analyze the role of dispersion in price rigidity on sectoral Phillips curves. They document that the average of the slopes of sectoral Phillips curves with heterogeneous price rigidity is larger than an economy with the average price rigidity.

product of labor is decreasing in output. Therefore, the firm sets a lower p_t^* than it would otherwise do in the absence of this endogenous feedback effect. The market power channel that we emphasize interacts with this endogenous feedback effect. In particular, the latter is amplified in a more competitive market (bigger ε). Because for the same amount of the price differential $(p_t^* - p_{t+k})$, the quantity differential is larger in a more competitive market. As a result, firms' prices respond less to shifts in marginal costs. In other words, the slope of the Phillips Curve is flatter. It follows that the degree of monetary non-neutrality is increasing in ε.

Although a multi-sector version of the model is not a simple weighted average of the multiple basic models, the intuition provided in this subsection does carry over. The remainder of this section will illustrate the quantitative importance of the outlined effect in a multi-sector model.

2.4 implications of dispersed markups

We assess the implications of the facts we documented in section (2.2) for the conduct of monetary policy based on a multi-sector NK model (Woodford, 2011, Carvalho, 2006). In this section, we discuss the implications for the degree of monetary non-neutrality and the Phillips Multiplier. We analyze the inflation targeting policy in section 2.5.

2.4.1 *The Multi-sector New Keynesian Model*

In this section, we present a dynamic multi-sector model (Woodford (2011), Carvalho, 2006) with heterogeneous degrees of market power, nominal rigidities, and sizes across sectors. The heterogeneity in market power is modeled in the following way. We assume that within a sector k, firm-level goods are aggregated into the sectoral aggregate goods C_{kt}

according to the following CES function:

$$C_{kt} \equiv \left[n_k^{-1/\varepsilon_k} \int_0^{n_k} C_{kt}(i)^{(\varepsilon_k - 1)/\varepsilon_k} di \right]^{\varepsilon_k/(\varepsilon_k - 1)}, \tag{2.9}$$

with an elasticity of substitution ε_k (hence the market power) that varies across sectors. The multi-sector economy is populated by a continuum of (0,1) of households, a fraction n_k of monopolistic competitive firms in sector k for $k = 1, 2.., K$, a government, and a central bank. A fraction of $1 - \theta_k$ of firms in sector k is allowed to reset their prices in each period. The degree of market power in each sector k is characterized by the elasticity of substitution across goods within the sector. Households consume the composite goods, buy a one-period risk-free government bond, supply labor to the sectoral competitive labor market, receive dividends (profits) from firms, and pay taxes or receive transfers from the government. Firms demand labor to produce and sell goods to households. The government issues government bonds, collects lump-sum taxes (or pays transfers) from (to) households. The central bank follows a Taylor rule. We leave the detailed description of the model to the Appendix 2.7.

Sectoral Phillips Curves. By solving the firms' optimization problem we obtain the New Keynesian Phillips Curve (NKPC) for each sector k:

$$\pi_{kt} = \kappa_k \widetilde{y}_t + \gamma_k \widetilde{y}_{R,kt} + \beta \mathbb{E}_t \pi_{k,t+1} \tag{2.10}$$

where $\widehat{y}_{R,kt} \equiv \widehat{y}_{kt} - \widehat{y}_t$, $\kappa_k \equiv \lambda_k(\sigma + \frac{\varphi+\alpha}{1-\alpha})$, $\gamma_k \equiv \lambda_k(\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})$, $\lambda_k \equiv \frac{(1-\beta\theta_k)(1-\theta_k)}{\theta_k}\Theta_k$, $\Theta_k \equiv \frac{1-\alpha}{1-\alpha+\alpha\varepsilon_k}$. Sectoral heterogeneities give rise to relative price (or quantity) dispersion across sectors; therefore, a full stabilization of both inflation and output gap is no longer feasible.[6] More

[6]An exception is when all relative output gaps are zero, e.g., in the presence of shocks to which all firms respond homogeneously. Here, instead, we consider sectoral shocks, and the firms' responses will be heterogeneous.

specifically, the aggregate Phillips Curve can be obtained by summing up the Sectoral NKPCs:

$$\pi_t = \sum_{k=1}^{K} n_k \kappa_k \widetilde{y}_t + \sum_{k=1}^{K} n_k \gamma_k \widetilde{y}_{R,kt} + \beta \mathbb{E}_t \pi_{t+1}, \tag{2.11}$$

where $\pi_t \equiv \sum_{k=1}^{K} n_k \pi_{k,t}$ denotes the aggregate inflation index (PCE).

Monetary Policy. The central bank sets the nominal interest rate according to a Taylor rule that targets an inflation index (π_t^{cb}) and the output gap:

$$i_t = \phi_\pi \pi_t^{cb} + \phi_y \widehat{y}_t + v_t, \tag{2.12}$$

where v_t denotes exogenous monetary policy shocks following an AR(1) process with persistence ρ_v. Before moving to the discussion of the optimal choice of π_t^{cb}, we assume that the central bank targets the PCE index π_t.

Calibration. We use the evidence from the empirical section to calibrate the parameters of the multi-sector model that govern sectoral heterogeneity. For the other parameters that characterize aggregate dynamics, we mostly follow Gali (2015). The model is calibrated to match the main categories that underly the PCE price index in the NIPA tables. Additionally, we decompose utilities into two categories to differentiate between core and non-core components. The model is calibrated at a quarterly frequency.

Homogenous Parameters. Most parameters are calibrated to values that are frequently used in the literature. We set the discount factor to 0.99, implying an (annualized) steady-state interest rate of 4%. $\sigma = 2$ implies that inter-temporal elasticity of substitution equal to 0.5. The Frisch elasticity of labor supply ($1/\varphi$) is set to be 1/5. The production function has decreasing returns to scale with $\alpha = 1/3$, a value commonly used in business cycle literature. We adopt interest rule coefficients suggested by Taylor (1993) as $\phi_\pi = 1.5$ and $\phi_y = 0.125$. Shock persistence are set to 0.8 and the variance of sectoral technology shocks and the monetary policy shock are chosen to be $\sigma_{ak} = 0.033$ and $\sigma_m = 0.044$ following Billi and Gali (2020). One important parameter is the across-sector elasticity of substitution, η. In one-sector models, there is only one parameter that governs aggregate markups. Here, we differentiate between within-sector elasticities, which finally govern sector-specific markups, and across-sector elasticities of substitution. We follow Atalay (2017) who calibrates the across sector elasticity using different approaches, finding that it is usually very small, implying that goods are rather inelastic across sectors. We follow his medium estimate and use $\eta = 0.5$.

Table 2.1: Calibration Homogenous Parameters

Utility function	$\sigma = 2$, $\varphi = 5$
Discount factor	$\beta = 0.99$
Production function	$\alpha = 1/3$
Technology shocks	$\rho_k = 0.8$, $\sigma^2_{ak} = 0.033$
Demand shock	$\rho_m = 0.8$, $\sigma^2_m = 0.044$
Elasticity of substitution	$\eta = 0.5$

Heterogeneous Parameters. Table 2.2 shows the calibrated values for all heterogeneous parameters. First, the heterogeneity in the size of a sector, n_k, corresponds to the steady-state share of expenditure the consumer assigns to this sector. Accordingly, we use the average personal consumption expenditures of households attributed to this sector over different decades.

Second, we calibrate the frequency of price adjustment. In the Calvo model, the frequency of price adjustment directly matches into price rigidity, since every quarter a fraction $(1 - FPA_k)$ of goods within the sector cannot adjust prices. We calculate the sales-weighted median of the frequency of price adjustment within each category as the respective measure of price stickiness. We convert monthly frequencies from the table into quarterly frequency. Here, there are two possible approaches: from the consumer or the producer side. While the former uses Nakamura and Steinsson (2008) estimates aggregated by Eusepi et al. (2011), the latter represents our aggregations based on the data on FPAs from Pasten et al. (2019).[7] The weighted average quarterly aggregate calibrated price rigidity – a fraction of firms that cannot change prices – in the 17-sector model is 0.6, which is in line with the aggregate rigidity (0.63) found in Gorodnichenko and Weber (2016a). One crucial observation is the similarity of many frequencies across these highly different approaches.[8] This is more surprising given that they do not share the same database as the origin. Instead, one is aggregated from ELI goods prices and the other from PPI prices. This gives us confidence for the aggregation of the markups that follow the same methodology as for the PPI frequencies.

Third, we calculate markups across sectors using the cost-shares approach at the firm level. In a first step, we aggregate firm-level markups at the three-digit NAICS level using Compustat declarations and take average values across decades. We then use PCE bridge tables to assign the NAICS sectors to the personal consumption categories. We use weights based on the producer value of goods. The markups we derive following this approach are reported in Table 2.2 and can be matched into sector-specific elasticities of substitution, ε_k.

[7]We are grateful to Michael Weber for sharing this data with us.

[8]Particularly interesting is the health care sector. The frequency based on the health care ELIs included in the CPI research database studied by Nakamura and Steinsson (2008) is 3.4 percent, implying an average duration of prices of 29 months. Given this high rigidity, Eusepi et al. (2011) deviate and choose an ad-hoc value of 8.3 to have an implied duration of 12 months to match spikes in October and January of the underlying data. Our estimates based on the PPI data generate a rigidity of 7.96, close to the implied duration of 12 months, supporting the view that prices in this sector change at least once a year.

Table 2.2: Calibrated Parameter Values

Sector name	Core	PCE share 1960s	PCE share 2000s	FPA PPI	FPA CPI	Markup 1960s	Markup 2000s
i) One sector							
PCE		100	100	14.17	12.93	1.3	1.6
ii) 2 sectors							
Manufacturing		51.67	34.78	11,57		1.2	1.58
Services		48.33	65.22	8,47		1.6	1.81
iii) 17 sectors							
Motor vehicles	X	6.51	5.03	38.66	36.6	1.27	1.30
Furnishings and household	X	4.93	3.26	10.76	9.2	1.50	1.94
Recreational goods	X	2.51	3.75	9.48		1.64	3.03
Other durable goods	X	1.45	1.80	6.55	10.9	1.39	1.65
Food (off-premises)		17.98	8.62	16.96	13	1.42	1.93
Clothing and footwear	X	7.97	3.94	6.92	32.2	1.41	2.41
Gasoline & energy goods		4.62	3.40	79.83	87.6	1.50	1.32
Other nondurable goods	X	8.90	8.90	15.59	10.4	1.41	1.73
Housing	X	15.17	17.09	29.79	10.3	1.52	1.90
HH Utilities Core	X	0.48	0.81	17.86	11.4	1.35	1.60
HH Utilities Non Core		2.62	2.37	33.82	38.5	1.72	1.33
Health care	X	6.31	16.78	7.96	8.3	1.49	1.20
Transportation	X	2.98	3.68	10.55	71.5	1.28	1.27
Recreation services	X	2.19	4.23	5.06	10	1.47	1.91
Food services	X	6.10	5.83	25.51		1.34	1.27
Accommodations	X	0.43	0.95	21.28		1.25	1.25
Other services	X	8.84	9.56	4.49	7.5	1.38	1.95

Note: This table shows the calibrated parameters of the different sector economies. Share and frequencies are in percentage points. FPA PPI is from Pasten et al. (2019) and FPA CPI is from Nakamura and Steinsson (2008). Markups are calculated using Compustat at the firm level and aggregated to PCE sectors using PCE bridge tables.

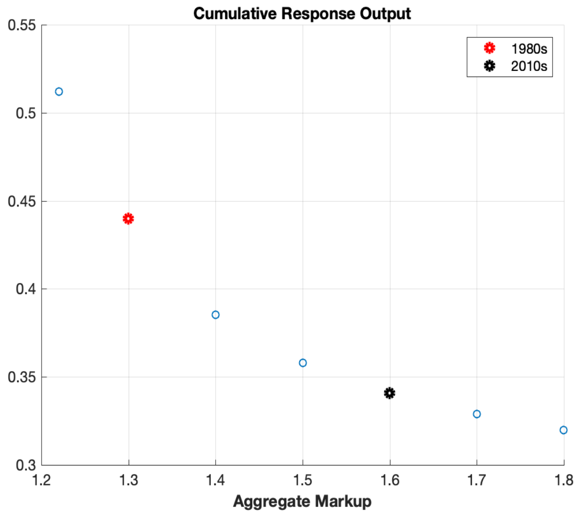
2.4.2 *Results: Money Non-Neutrality*

This section reports the first theoretical results on the relationship between non-neutrality and the Phillips multiplier with different components of the markup distribution – (i) mean and (ii) dispersion. Note that we consider mean and dispersion as stand-ins for the whole distribution of markups. In this sense, it is likely that higher moments like kurtosis will also affect non-neutrality. While the analysis of those exceeds this paper's scope, the aim is to motivate to look at the whole distribution instead of solely the average markup to study the relationship between non-neutrality and markups. We do so in a simple three-sector version of the multi-sector model, where we will control the moments of the markup distribution. The homogenous parameters follow the exposition in the last section, and sector sizes and the frequency of price adjustment are calibrated to their average levels. In the second part, we will then look at predictions concerning the Phillips multiplier in the different decades of the model's seventeen-sector calibration as outlined in Table 2.2.

Non-Neutrality Decreases in Average Markup. First, we study the effect of increases to the average markup on non-neutrality in the multi-sector model while keeping other moments of the distribution constant. This resembles the exercise performed in the motivational example with the difference that there we showed the effect of an increase in a one-sector model (as, e.g., in Gali, 2015). In particular, we calibrate different versions of the model that feature heterogeneity in sectoral markups and differ in the average markup across versions.

Figure 2.5 shows the cumulative response of output to a 25bp expansionary monetary policy shock in different calibrations of the model with the average markup of those calibrations on the horizontal dimension. We find a clear negative relationship between markups and non-neutrality as in the text-book one-sector model. With less compe-

FIGURE 2.5: Monetary Non-neutrality and Aggregate Markup

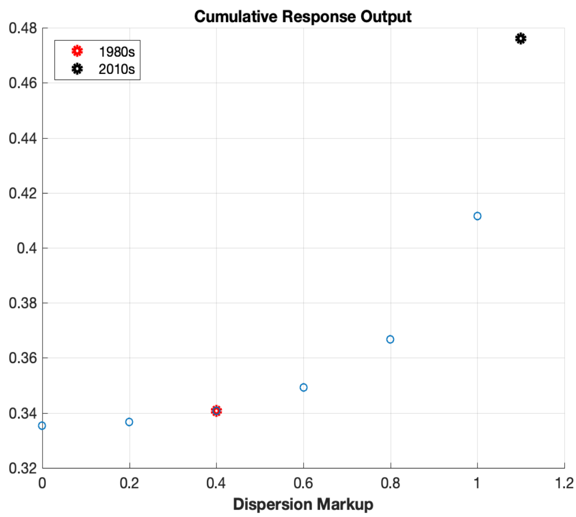


Note: This figure shows the cumulative output response to a 25bp expansionary monetary policy shock in different three-sector calibrations of the multi-sector model that differ in the aggregate degree of market power.

tition in an economy, monetary policy becomes less effective. We also highlight the size of average markups in the economy in the 1980s and 2000s. According to the model, the increase in markups of 30 percentage points decreased non-neutrality by over 20 percent. This is a direct application of Proposition 2.3.1 since an increase in the average markup is related to a decrease in the aggregate elasticity of substitution, ε_k. Intuitively, if demand is less elastic to changes in prices, demand increases will lead to larger changes in prices with smaller output adjustments. Consequently, money non-neutrality becomes smaller for all sectors. This also implies that the Phillips curve multiplier becomes steeper for all sectors and, thus, also on the aggregate. In conclusion, the results in the multi-sector model are consistent with the one-sector model.

Non-Neutrality Increases in Markup Dispersion. As we have seen in the empirical evidence, dispersion measured by the standard deviation of idiosyncratic markups more than doubled over time. Therefore, we are interested in measuring the effect of increases in the dispersion of sectoral markups on non-neutrality in the multi-sector model. For this, we consider the three-sector model again, but this time we keep the average markup constant. Instead, we increase step-wise markup dispersion by decreasing (increasing) the lower (higher) markups in the same proportions. In the one-sector model, a change in the composition of sectoral markups does not change any predictions as long as the aggregate markup stays the same.

Figure 2.6: Monetary Non-neutrality Increases in Markup Dispersion



Note: This figure shows the cumulative output response to a 25bp expansionary monetary policy shock in different three-sector calibrations of the multi-sector model that differ in the dispersion of markups but have the same aggregate degree of market power.

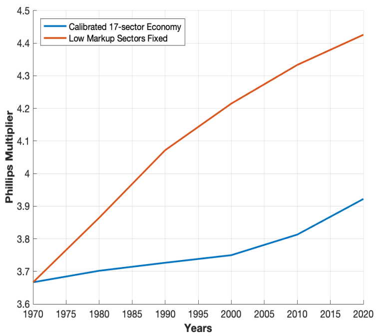
Figure 2.6 shows the cumulative response of output to a 25bp expansionary monetary policy shock when the dispersion of markups across sectors increases (horizontal axis). Figure 2.6 shows a clear positive relationship between markup dispersion and non-neutrality. With more dispersed markups in an economy, monetary policy becomes more effective. Specifically, the increase in dispersion in the U.S. economy from the 1960s until today could have increased non-neutrality by more than 40 percent. The reason is the non-linearity in the size of the output response to the level of markups outlined in Proposition 2.3.1. Intuitively, decreasing the markups of the low markup sectors increases their non-neutrality. This increase is larger than the decrease in non-neutrality due to the equivalent increase in the markups of the high markup sectors. Consequently, the increase in dispersion over time could offset the predictions of increasing markups in the one-sector and multi-sector models. We will study whether this is indeed the case in the next section, where we calibrate the multi-sector model to the actual moments we observed for the U.S. between 1967 - 2017.

Phillips Multiplier in the U.S. Now, we consider the seventeen-sector version of the multi-sector model and calibrate it to the data described in the previous section and summarized in Table 2.2. To identify the effects of changes in markup distribution, we fix sectoral sizes at their 2000 level. Figure 2.7 shows the Phillips multiplier for each decade from the 1960s until the 2010s. Our analysis defines the Phillips multiplier as the ratio between the cumulative responses of inflation and output to a 25bp expansionary monetary policy shock. We find a relatively stable Phillips multiplier (blue line) and monetary non-neutrality (blue line in the top panel of Panel B) over time.[9] The reason for these stable statistics is the spreading-out of the markup distribution. The effects of a decline in the left tail offset the effects of an increase in the markup distribution's right tail. This is verified by

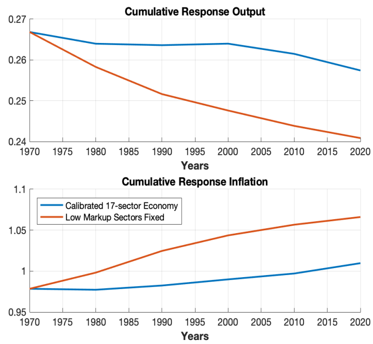
[9]Due to aggregation effects, the Phillips multiplier and slope of the Phillips curve tend to be larger in multi-sector models than in a one-sector model. The reason is that κ_k in equation (2.10) is non-linear in θ_k and ε_k (see Imbs et al. (2011)).

the red lines in Panel A and B: had the left tail remained constant over time – in the counterfactual, we fix the markups of the four sectors with the lowest markups at their 1960s value – the Phillips multiplier would have increased.

Figure 2.7: The Implied Evolution of the Phillips Multiplier in the U.S.: The Role of the Distribution of Markups



(a) Phillips Multiplier



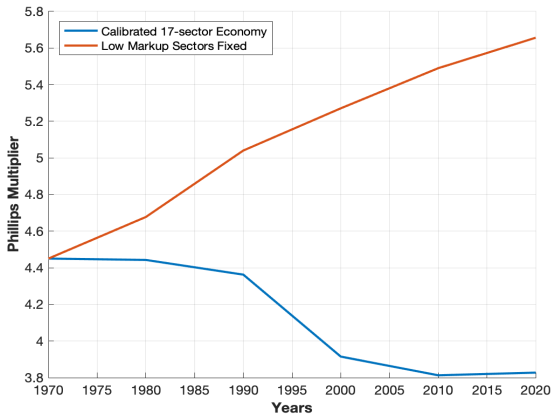
(b) Cumulative Responses

Note: Panel (a) of this Figure shows the dynamic multiplier (blue line) as the ratio of the cumulative responses of inflation and output to a 25bp expansionary monetary policy shock in seventeen-sectors calibrations of the multi-sector model keeping the sizes of sectors fixed at their 1960s value over time. Panel (b) shows each cumulative response separately. In the counterfactual (red line), we fix the markups of the four sectors with the lowest markups at their 1960s value.

Next, we calibrate our seventeen-sector over time to also consider the changes in sector sizes, n_k. Figure 2.8 plots the results. Due to changes in sector sizes, the degree of monetary non-neutrality in the U.S. has increased over time. The inflation-output tradeoff summarized by the Phillips multiplier has declined by roughly 20%. Moreover, Panel A of Figure 2.8 shows another important result. The graph compares the calibrated 17-sectors model with a counterfactual that keeps the lowest markups constant. The prediction on the change in the

Phillips multiplier would otherwise be of opposite sign and about 50% higher. This confirms the importance of monitoring the whole markup distribution.

FIGURE 2.8: The Implied Evolution of the Phillips Multiplier in the U.S.: The Role of Changes in Sizes

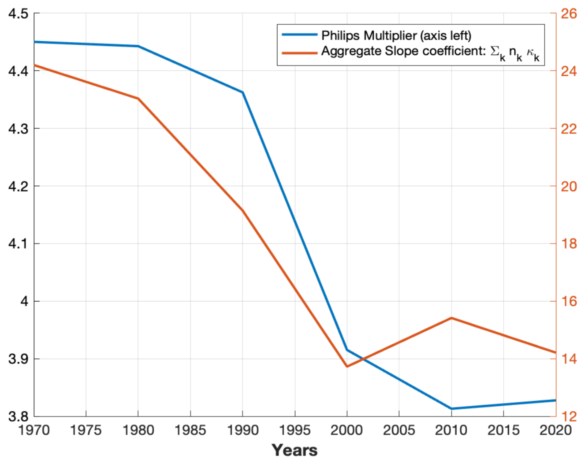


Note: This Figure shows the dynamic multiplier (blue line) as the ratio of the cumulative responses of inflation and output to a 25bp expansionary monetary policy shock in the seventeen-sectors calibration of the multi-sector model with changes in the sizes of sectors over time. In the counterfactual (red line), we fix the markups of the four sectors with the lowest markups at their 1960s value.

Next, we want to understand where the decreasing Phillips multiplier in Panel A of Figure 2.8 comes from. To understand these results, we look at the aggregate New Keynesian Phillips curve, equation (2.11). The slope of the Phillips curve is the elasticity of inflation to changes in the output gap. In the present case, it is represented by the coefficient in front of the output gap, $\bar{\kappa} = \sum n_k \kappa_k$.[10]

[10] Chapter one shows that in multi-sector models, the slope of the Phillips curve is measured with a bias since relative output gaps are possibly correlated with the aggregate output gap. In the present case, the bias appears not to be driving the results and of small size.

Figure 2.9: Phillips Multiplier vs. Phillips Curve



Note: This Figure shows the estimated Phillips multiplier from Figure 2.8 together with the aggregate slope coefficient – defined as the weighted sum of sectoral slope coefficients, κ_k – over time.

Figure 2.9 shows that the decrease in the Phillips multiplier (solid blue line) is almost entirely explained by the decreasing aggregate slope coefficient (solid red line) in the model. The slope coefficient, $\bar{\kappa}$, depends non-linearly on the sectoral elasticity of substitution. Consequently, changes to smaller ε_k have a smaller impact on the aggregate coefficient than those to larger ε_k. This is confirmed if we look at the cumulative responses of output and inflation in Figure 2.9. The Phillips multiplier is smaller because the slope coefficient decreases; larger output fluctuations require smaller deviations of inflation (blue line).

2.5 The Optimal Inflation Target Policy

Welfare Loss Function Before moving to the central bank's problem, we will derive the welfare loss function, which is the objective of the central bank. Following Rotemberg and Woodford (1997, 1999) and

Woodford (2002), we derive the welfare loss function as the second-order approximation of the representative consumer's period welfare loss expressed in consumption equivalent variation (CEV):

$$L = \sum_{k=1}^{K} \frac{\varepsilon_k}{\lambda_k} n_k var(\pi_{kt}) + (\sigma + \frac{\varphi + \alpha}{1-\alpha}) var(\widetilde{y}_t) \tag{2.13}$$
$$+ (\eta^{-1} + \frac{\varphi + \alpha}{1-\alpha}) \sum_{k=1}^{K} n_k var(\widetilde{y}_{R,kt}),$$

where $\lambda_k \equiv \frac{(1-\beta\theta_k)(1-\theta_k)}{\theta_k}\Theta_k$ defined as above. Normalize the weights on π_{kt} such that $\sum \phi_k = 1$:

$$L = \sum_{k=1}^{K} \phi_k var(\pi_{kt}) + \lambda_y var(\widetilde{y}_t) + \lambda_{R_y} \sum_{k=1}^{K} n_k var(\widetilde{y}_{R,kt}), \tag{2.14}$$

where

$$\phi_k = \frac{n_k \varepsilon_k \lambda}{\lambda_k}, \qquad \lambda_y = (\sigma + \frac{\varphi+\alpha}{1-\alpha})\lambda, \qquad \lambda_{Ry} = (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})\lambda,$$

and λ is defined as:

$$\lambda \equiv (\sum_{0}^{K} n_k \varepsilon_k \lambda_k^{-1})^{-1}.$$

See the Appendix 2.7 for details of the derivations.

Interestingly, by allowing for sectoral heterogeneity in market power, inflation of a sector with a higher elasticity of demand enters in the welfare loss function with a bigger relative weight, i.e., $\frac{\partial \phi_k}{\partial \varepsilon_k} > 0$.

The Optimal Inflation Index Stabilization Policy The central bank adopts inflation targeting as the means of conducting monetary policy. This is the case for many central banks around the world. We assume that the target rate is zero (the steady-state inflation rate), and the goal is always achieved. This is equivalent to a Taylor rule with strict inflation index targeting. The monetary instrument is the ex-ante choice of an inflation index that the central bank stabilizes ex-post. This question can be formulated as the following:

$$\min_{\{\omega_k\}} L = \min_{\{\omega_k\}} \sum_{k=1}^{K} \phi_k var(\pi_{kt}) + \lambda_y var(\widetilde{y}_t) + \lambda_{R_y} \sum_{k=1}^{K} n_k var(\widetilde{y}_{R,kt}), \tag{2.15}$$

subject to equilibrium conditions, resources constraints, and $\sum_{k=1}^{K} \omega_k \pi_{kt} = 0$.

Previous studies have uncovered two main results. First, if sectors share the same degrees of nominal rigidities and market competition, the stabilization of PCE is optimal. Second, it is optimal to give higher weight to the sector with a higher degree of nominal rigidity. The remaining of this paper is to investigate the role of market power and, in particular, how it might interact with the stickiness channel.

2.5.1 *Special Cases*

We begin with analyzing a limiting case in which one sector is infinitely close to perfect competition[11], i.e., $\varepsilon_k \to \infty$. In this case, the loss function collapses to:

$$L \to var(\pi_k).$$

It follows immediately that:

Proposition 2.5.1 *In the limiting case $\varepsilon_k \to \infty$ and $\theta_k \neq 0$, the optimal monetary policy is to set $\pi_k = 0$.*

This does not mean that the welfare loss under the optimal monetary policy is zero. In fact, due to asymmetric shocks, the aggregate, the

[11] It only makes sense to talk about the infinitely close case because in the limiting case with perfect competition, firms are price takers. Therefore the firm's problem discussed in the previous section, price setter firms, would not carry over.

relative output gap, and inflation in the remaining sectors fluctuate, which gives rise to welfare losses. It means that if goods in sector k are almost perfect substitutes, then, in terms of welfare loss, stabilizing inflation in this sector is infinitely more valuable than stabilizing any other variables. This is the case because, with a flat demand curve and nominal rigidity, price dispersion that arises from inflation leads to an infinitely big dispersion in output.

Next, we investigate whether the competition channel affects the optimality of core inflation stabilization suggested by Aoki (2001) and Benigno (2004).

Proposition 2.5.2 *If the price is flexible in sector k, independent of the relative market power, the optimal weight for this sector is zero.*

Proof: see Benigno (2004).
If the price is flexible, inflation does not lead to price dispersion. Therefore welfare loss originating from inflation is trivial no matter how competitive the market is. A more interesting interaction between market power and nominal rigidity arises in the general case.

2.5.2 *Inflation Targeting Policies in a More Aggregated Two-Sector Model*

We begin by illustrating the mechanism based on a two-sector model calibrated to the manufacturing and service sector in the U.S. Those two sectors represent a significant fraction of aggregate production in the U.S. Therefore, the results we find below represent the findings for a relatively (compared to a 17-sector model) more aggregated model.

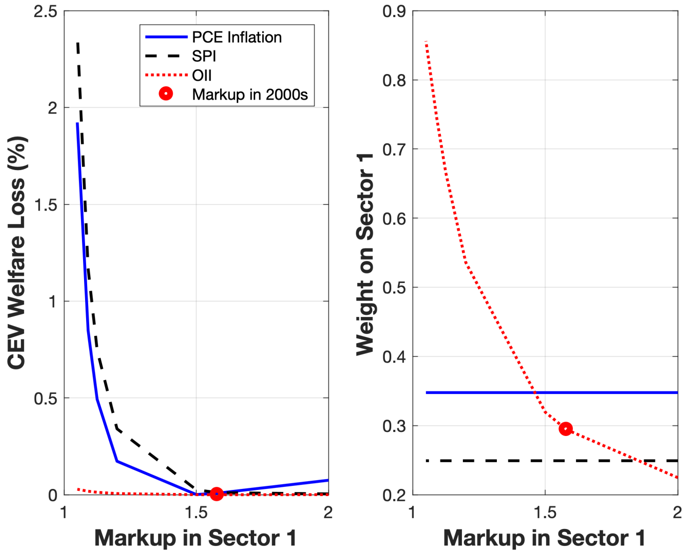
Calibration. Unless otherwise specified, the model's parameters are calibrated to be those reported in Table (2.1) and (2.2) in the 2000s. The heterogeneous parameters are calibrated to match their counterparts in the manufacturing (sector 1) and service (sector 2) sectors in the U.S.

from 2000-2010. The sectoral degrees of nominal rigidity are $\theta_1 = 0.69$ and $\theta_2 = 0.77$ for the manufacturing and service sectors. They are chosen to match the monthly frequency of price adjustments based on the PPI and are similar to those reported in Gorodnichenko and Weber (2016b). The sectoral elasticities of substitution are calibrated to be $\varepsilon_1 = 2.72$ and $\varepsilon_2 = 2.25$ in order to match markups in manufacturing (1.58) and service sectors (1.81) estimated by DeLoecker et al. (2020). Those markups are higher than the values that are typically assumed in the literature: 1.1 or 1.2. We provide a robustness check using those commonly used values, and qualitatively the results are unchanged. What matters is the markup in the service sector is higher than in manufacturing, which is confirmed by Christopoulou and Vermeulen (2012) in their estimates of markups for both the U.S. and the Euro Area.

The Competition Channel Offsets the Stickiness Channel. We conduct a welfare analysis under alternative inflation index stabilization policies: the optimal inflation index (OII), the stabilization of the PCE inflation index, and the stickiness-based index (SPI). The optimal weights of the policies are chosen to minimize the welfare loss. To compute the welfare minimizing weights of the SPI, we consider a model with only heterogeneities in price rigidities as in Aoki (2001) or Benigno (2004). For the optimal inflation index, we consider a model with heterogeneities in both price rigidities and markups. We then compare the welfare implications in an economy with both heterogeneities, where we consider different values of the steady-state markup in sector one. Figure 2.10 reports the results. The reported welfare loss is the CEV defined above in deviation from the CEV under the optimal monetary policy. The left panel shows the welfare loss under alternative policy rules, and the right panel plots the associated weights.

Interesting results arise when comparing the performance of PCE stabilization with the stabilization of SPI. When markup in sector 1

FIGURE 2.10: Inflation Targeting Policy in a Two-Sector Model



Note: Welfare loss as a function of markup in sector 1 under alternative policies. The welfare loss is the corresponding CEV in deviation from the CEV under optimal monetary policy. The red dot corresponds to the point where the markup in sector 1 equals 1.58 – the markup for the manufacturing sector in the data.

is large enough, stabilizing the inflation index based on stickiness as recommended by Benigno (2004) and Mankiw and Reis (2003) is welfare improving compared to the stabilization of PCE. However, if sector 1 is competitive with a small markup, stabilizing PCE is superior. Hence, stabilizing SPI is a sensible policy advice (compared to the stabilization of PCE) if the sector with a relatively more flexible price (sector 1) is associated with a bigger markup. This is not the case in the data: Figure(2.2) shows a negative relationship between price flexibilities and steady-state markups.[12] This suggests that the competition channel works against the stickiness channel. The red dot in Figure 2.10 points

[12]Costly price adjustment models developed by Barro (1972), Sheshinski and Weiss (1977) and Golosov and Lucas (2007) predict more flexible price in a sector with higher competition.

out the steady-state markup in sector 1 in the data. Interestingly, in the calibrated two-sector model that we consider here, the competition channel fully offsets the stickiness channel. As a result, the stabilization of PCE or SPI leads to welfare losses of a similar amount.

2.5.3 *Inflation Targeting Policies in a Calibrated Seventeen-Sector Model*

We now consider the inflation targeting policies in a seventeen-sector model. The model's parameters are calibrated to be those reported in Table (2.1) and (2.2) for the 1960s and the 2000s subsamples. Sectors are heterogeneous in their degree of nominal rigidities, market powers, and sector sizes. We consider four inflation index weights: (i) the optimal inflation index (OII) weights computed using a model with the three heterogeneities; (ii) the stickiness-based inflation index (SPI) weights are computed without the consideration of heterogeneity in market power, i.e., sectoral markups are set to the economy-wide average level; (iii) the markup weights are inferred from an economy where there is no heterogeneity in nominal rigidities, and all sectoral frequencies of price adjustment are set to the economy-wide average level; (iv) the PCE weights are given by the size of a sector, inferred from their expenditure share in the data, n_k.

The Importance of the Markup Channel Increased Over Time. Table (2.3) reports the percentage gain in the welfare loss under alternative inflation targeting policies relative to that under the PCE stabilization policy. That is, we report $100 * (L_{PCE} - L_i)/(L_{PCE})$, where $i = \{OII, SPI, Markup\}$ and L_i denotes the welfare loss under an inflation targeting policy i. The role of the market channel played a minor role in the 1960s. However, in the twenty-first century, the market channel has gained importance in the design of inflation targeting policy. The stabilization of SPI leads to a welfare loss that is 11.3% smaller than the loss under the stabilization of the PCE. In contrast, the OII that takes market power heterogeneity into consideration leads to a welfare gain of 17.4%.

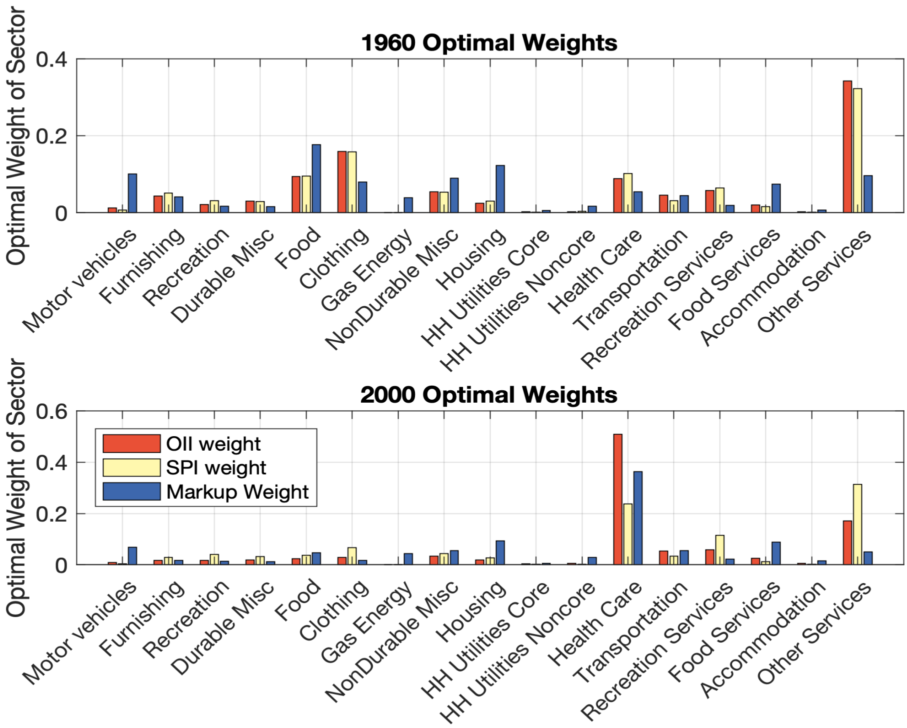
Table 2.3: Welfare Gain Relative to the Stabilization of PCE

Inflation Targeting Policies	1960s	2000s
OII	7.5%	17.4%
SPI	7.4%	11.3%
Markup	0.3%	8.2%

The pattern on the importance of the markup channel over time is confirmed in Figure 2.11, where we plot the OII weights, SPI weights, and the markup weights for each sector in the two subsamples. A clear pattern stands out: in the 1960s, the SPI weights almost coincide with the OII weights, whereas the difference between the two weights is more pronounced in the 2000s. The service sectors are plotted towards the end of the x-axis. The optimal weights for those sectors in the 2000s (red bar) lay in between the SPI weights (yellow bar) and the markup weights (blue bar), indicating that the markup channel partly offsets the stickiness channel as we explained in the two-sector calibration. However, the major change between the 1960s and the 2000s is the weight on the health care sector. Our OII weight on the health care sector is of a similar magnitude as in Eusepi et al. (2011). The health care sector deserves a significant weight in the 2000s because it is a sector with a large size, low markup, and low frequency of price adjustment. As shown in Figure 2.2, the health care sector is an outlier in the markup and price flexibility relationship.

The Evolution of the OII. We construct the time series of realized OII over time by combining the OII weights with realized sectoral price

Figure 2.11: Optimal Inflation Index Weights (OII) 1960s vs. 2000s.



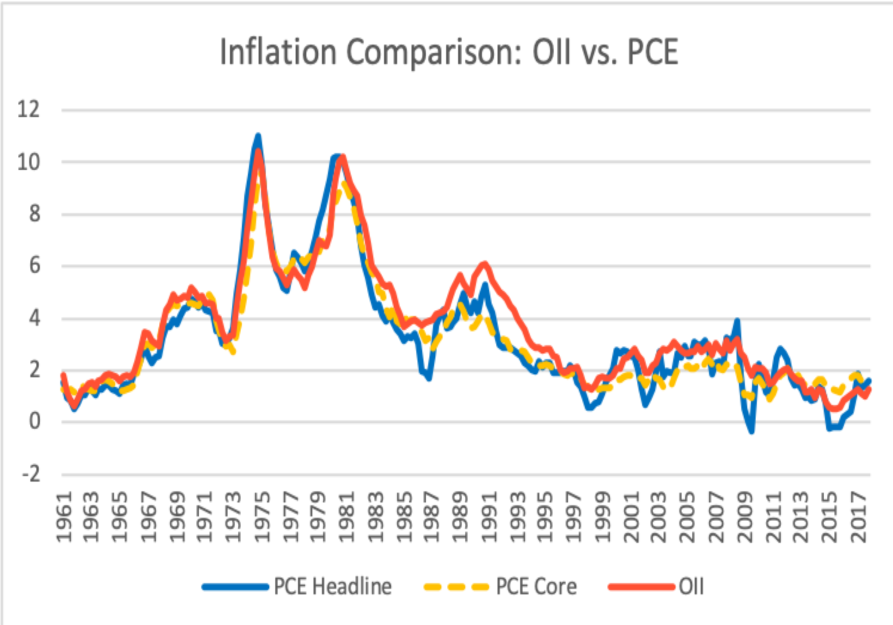
Note: This figure compares the sectoral weights for each policy considered: (i) the OII weights (red), (ii) the SPI weights (yellow) and (iii) the markup weights (blue). They are computed from the 17-sectors version of the model calibrated either to the heterogeneous parameters in the 1960s (first panel) or to those in the 2000s (second panel).

dynamics. We build the OII as

$$\pi_t^{OII} = \sum_{k=1}^{K} \omega_{k,t}^{OII} \pi_{k,t},$$

where the weights $\omega_{k,t}^{OII}$ are the optimal weights derived in the 17-sectors calibration in the multi-sector model with heterogeneity in markups and nominal rigidity in each decade. The sectoral inflation rates, $\pi_{k,t}$,

FIGURE 2.12: The Optimal Inflation Index



Note: This figure compares the OII (yellow dashed line) with the historical series of PCE headline (blue solid line) and PCE core (red solid line) inflation. The OII is constructed by combining the OII weights from each decade, $\omega_{k,t}^{OII}$, with the historical time series of the inflation rates of the PCE components, $\pi_{k,t}$. For consistency, the PCE inflation rates are calculated by averaging the sectoral weights for each decades instead of using yearly data on weights.

are the historical time series of realized inflation rates in the 17 PCE categories. Replacing the optimal weights by the expenditure shares of sectors, n_k, one can derive the time series of realized PCE inflation.

Figure 2.12 plots the OII together with the headline and the core PCE. During the Great Moderation periods, the OII was consistently higher than the two PCE measures that the Fed relies on in their policy analysis. This demonstrates that the OII stabilization cannot be achieved by monitoring a weighted average of the headline and the core PCE. During the periods following the Great Recession of 2008, similar to other inflation measures, the OII is below the 2% target.

2.6 CONCLUSION

We have witnessed a substantial change in the average markups in the U.S. in the last decades. In this paper, we document that those changes are heterogeneous across sectors. Particularly, while there is a persistent increase in the right tail of the markup distribution, firms in the left tail suffered a persistent decline in the steady-state markups. This finding has important implications for the conduct of monetary policy.

First, the degree of monetary non-neutrality is higher in a model that features heterogeneity in steady-state markups compared to a model with homogenous market powers.

Second, changes in the markup distribution have minimal impact on the Phillips Multiplier in the U.S. due to the offsetting effects of the increase in the right tail and the decrease in the left tail of markup distribution.

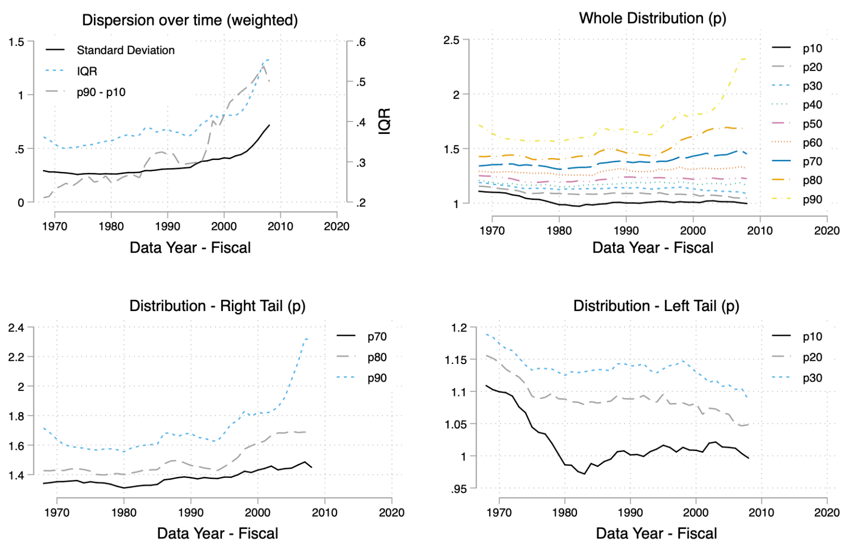
Third, addressing markup dispersions is relevant for the optimal inflation targeting policy. Based on a two-sector model calibrated to the manufacturing and service sectors in the U.S, we show that the heterogeneity in markups offsets the stickiness channel. This finding challenges the conventional wisdom among academics and policymakers that a sector's optimal weight in the OII is proportional to its relative price rigidity; instead, heterogeneities in market powers matter, too. Crucial to this finding is the negative relationship between sectoral steady-state markups and the frequency of price adjustment that we document in this paper. Moreover, we show, based on the model calibrated to seventeen sectors in the U.S., the importance of the markup channel became more important in the 2000s as compared to the 1960s.

2.7 APPENDIX

Empirical Results for Alternative Markups

We reproduce the empirical results concerning the markup distribution by calculating markups following the approach outlined in DeLoecker et al. (2020) using Compustat again as data source. The difference to the cost-share approach is that we calculate sectoral output elasticities, ε_{Q,X_i}, by estimating a four-digit industry-specific production function. Figure 2.13 shows that the results show the same patterns as our baseline estimates.

FIGURE 2.13: Steady State Markups by Quantiles and their Dispersion: Robustness



Note: This figure documents the dispersion and quantiles in markups calculated by the production approach and Compustat data following DeLoecker et al. (2020) from 1967-2017.

Estimation of Negative Correlation Between Markups and Frequency

We want to estimate the size and sign of the relationship between markups and nominal rigidity. To test the relationship, we observe matched markup-frequency pairs at the 3-digit industry level from our constructed dataset. Since frequencies have been calculated over the 2005-2011 period, we calculate sectoral markups over the same timespan.

Columns (1) and (4) of Table 2.4 document that firms with higher markups change prices less often. To control for possible omitted variable bias, we add two sets of control variables: (i) firm characteristics such as size, output, fixed capital stock, and (ii) firm- and time-specific volatilities. With the former fixed effects at the 2-digit industry level, we intend to control other sector-specific characteristics we did not consider. Columns (2) and (5) show that the estimated relationship is of the same sign, size, and significance, but the adjusted R-squared has increased. Firms could change prices more often when they are subject to more volatile shocks. In columns (3) and (6), we control for those effects via sector- and time-specific volatilities. Note that the number of observations decreases in this case because there were not sufficient observations to calculate volatilities for each sector. With these controls, the estimated correlation becomes stronger and remains of the same sign.

Table 2.4: Regression Markups and Nominal Rigdity 2005-2011

		FPA			Duration	
	(1)	(2)	(3)	(5)	(6)	(7)
Markup	-0.057***	-0.066***	-0.079***	2.278***	2.475***	2.836***
	(0.016)	(0.016)	(0.019)	(0.54)	(0.54)	(0.56)
Sales		0.0376*	-0.0732		-0.0687	5.07***
		(0.0157)	(0.045)		(0.541)	(1.35)
Employment		-0.134	0.999		-24.3	-76.9*
		(1.01)	(1.21)		(34.9)	(36.2)
Output		-0.374	-0.144		-4.38	-12.6
		(0.252)	(0.261)		(8.69)	(78.1)
Capital		0.954**	0.679*		-6.08	3.51
		(0.333)	(0.337)		(1.15)	(1.01)
$Volatility_i$			-0.0254			0.102
			(0.102)			(3.06)
$Volatility_t$			3.59*			-16.6***
			(1.39)			(4.16)
FE	NO	YES	YES	NO	YES	YES
N	501	501	377	501	501	377
adj. R^2	0.022	0.253	0.244	0.032	0.202	0.294

Note: Table shows results of regressing FPA (duration) on markups. Controls include sales, output (in million), employment and capital (per 100.000) and volatitility of sales across time or within sector. Fixed effects are at the 2-digit sector level. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

DESCRIPTION OF THE MODEL

Households

A representative household seeks to maximize the following utility function:

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left[\frac{C_t^{1-\sigma}}{1-\sigma}-\sum_{k=1}^{K}\frac{N_{k,t}^{1+\varphi}}{1+\varphi}\right],$$

subject to budget constraint:

$$P_tC_t+Q_tB_{t+1}\leq B_t+\sum_{k=1}^{K}W_{kt}N_{kt}+\sum_{k=1}^{K}T_{kt},$$

where P_t denotes the aggregate price defined below, Q_t denotes the price at time t of a one period bond that pays B_{t+1} at time $t+1$, W_{kt} the sectoral wage and T_{kt} the lump-sum transfer including profit from firms. There are K sectors in the economy, each of those sectors requires a sector-specific labor N_k. The aggregate consumption that enters utility function is a CES aggregate of K subindices:

$$C_t\equiv\left[\sum_{k=1}^{K}n_k^{1/\eta}C_{kt}^{(\eta-1)/\eta}\right]^{\eta/(\eta-1)},, \tag{2.16}$$

with the elasticity of substitution across sectors $\eta>0$, and n_k denotes the size of the sector k with $\sum_{k=1}^{K}n_k=1$. Each subindices C_{kt} is a CES aggregate of the following form:

$$C_{kt}\equiv\left[n_k^{-1/\varepsilon_k}\int_0^{n_k}C_{kt}(i)^{(\varepsilon_k-1)/\varepsilon_k}di\right]^{\varepsilon_k/(\varepsilon_k-1)}, \tag{2.17}$$

with an elasticity of substitution ε_k that varies across sectors.

The implied sectoral prices index are:

$$P_{kt}\equiv\left[n_k^{-1}\int_0^{n_k}p_{kt}(i)^{1-\varepsilon_k}di\right]^{1/(1-\varepsilon_k)}. \tag{2.18}$$

The implied aggregate price index is:

$$P \equiv \left[\sum_{k=1}^{K} n_k P_{kt}^{1-\eta} \right]^{1/(1-\eta)}. \tag{2.19}$$

Solving the consumers' problem regarding the optimal allocation of demand across varieties yields the following demand functions:

$$C_{kt}(i) = \frac{1}{n_k} C_{kt} \left(\frac{P_{kt}(i)}{P_{kt}} \right)^{-\varepsilon_k}, \qquad C_{kt} = n_k C_t. \left(\frac{P_{kt}}{P_t} \right)^{-\eta} \tag{2.20}$$

The former is the demand function faced by an individual firm i in sector k, and the one on the right is the sectoral demand faced by sector k. It is worth emphasizing that the price elasticity of demand faced by firm i in sector k is $-\varepsilon_k$, the same magnitude as the elasticity of substitution with the opposite sign (downward sloping). This is intuitive: the higher the elasticity of substitution, the easier it is for a consumer to substitute goods i by goods j in the same sector. Hence, the more elastic is the demand, the more competitive this sector is. In the limiting case of $\varepsilon_k \to \infty$, the market is perfectly competitive.

Firms

There are K sectors in the economy, with a continuum of monopolistic competitive firms operating in each of those sectors. All sectors share the production function of the same functional form but are subject to different shocks:

$$Y_{kt} = e^{a_{kt}} N_{kt}^{1-\alpha}. \tag{2.21}$$

Firms are subject to nominal rigidity *à la* Calvo (1983): Each firm may reset its price with probability $1 - \theta_k$. Hence, the log level price at sector k, p_{kt}, evolves as the following:

$$p_{kt} = \theta_k p_{k,t-1} + (1 - \theta_k) p_{kt}^*,$$

where p^*_{kt} is the optimal price that a reoptimizing firm at sector k would set. This is the solution to the following problem:

$$\max_{P^*_{kt}} \sum_{h=0}^{\infty} \theta_k^h \mathbb{E}_t \Big\{ Q_{t,t+h} \big(P^*_{kt} Y_{k,t+h|t} - \Psi_{k,t+h}(Y_{k,t+h|t}) \big) \Big\}, \tag{2.22}$$

subject to its demand constraints specified in (2.20). Here $Q_{t,t+h} \equiv \beta^k (C_{t+h}/C_t)^{-\sigma}(P_t/P_{t+h})$ denotes the stochastic discount factor, $\Psi_{k,t+h}$ denotes the cost function and $Y_{k,t+h|t}$ is the output for a firm in sector k that last reset its price in period t.

The optimality condition implied by the firm's problem is:

$$\sum_{h=0}^{\infty} \theta_k^h \mathbb{E}_t \Big\{ Q_{t,t+h} Y_{k,t+h|t} \big(P^*_{kt} - \frac{\varepsilon_k}{\varepsilon_k - 1} \Psi'_{k,t+h}(Y_{k,t+h|t}) \big) \Big\} = 0.$$

Thus, the desired markup, defined as the markup under flexible price, is equal to $\frac{\varepsilon_k}{\varepsilon_k - 1}$. The frictionless markup is decreasing in ε_k: The monopolistic competitive firm charges a lower markup in a more competitive market.

Equilibrium

Solve the household's problem and log-linearize to obtain the dynamic IS equation:

$$\widetilde{y}_t = \mathbb{E}\widetilde{y}_{t+1} - \frac{1}{\sigma}[i_t - \mathbb{E}\pi_{t+1} - r_t^N)], \tag{2.23}$$

where

$$\widetilde{y}_t \equiv y_t - y_t^N, \qquad y_t^N = \psi^a \sum_{k=1}^{K} n_k a_{kt}, \qquad r_t^N \equiv \sigma \psi^a \sum_{k=1}^{K} n_k \mathbb{E}_t \triangle a_{k,t+1},$$

with $\psi^a \equiv \frac{(1+\varphi)}{\sigma(1-\alpha)+\varphi+\alpha}$. Throughout this paper, a variable with tilde denotes this variable in deviation from its natural level and a variable with hat denotes this variable in deviation from its steady-state. Solving the firms' optimization problem and log-linearize, we obtain the New Keynesian Phillips Curve (NKPC) for each sector k:

$$\pi_{kt} = \lambda_k(\widehat{mc}_{kt} - \widehat{p}_{R,kt}) + \beta\mathbb{E}_t\pi_{k,t+1}, \tag{2.24}$$

where $\lambda_k \equiv \frac{(1-\beta\theta_k)(1-\theta_k)}{\theta_k}\Theta_k$, $\Theta_k \equiv \frac{1-\alpha}{1-\alpha+\alpha\varepsilon_k}$, $p_{R,kt}$ is the sector k's relative price (relative to aggregate price), defined as $p_{kt} - p_t$. And $\widehat{mc}_{kt}$ is the real marginal cost in sector k, which is defined as:

$$\widehat{mc}_{kt} = \sigma(\widehat{y}_t - \widehat{y}_t^N) + \frac{\alpha+\varphi}{1-\alpha}(\widehat{y}_{kt} - \widehat{y}_{kt}^N) + \eta^{-1}(\widehat{y}_t^N - \widehat{y}_{kt}^N). \tag{2.25}$$

In the derivations of $\widehat{mc}_{kt}$, we have used household's labor supply equations and the fact that $\widehat{mc}_{kt}^N = -\eta^{-1}(\widehat{y}_{kt}^N - \widehat{y}_t^N)$ as it is implied by the sectoral demand function together with firms' frictionless desired prices. Plug (2.25) into (2.24) and replace $p_{R,kt}$ by $-\eta^{-1}\widehat{y}_{R,kt}$, where $\widehat{y}_{R,kt} \equiv \widehat{y}_{kt} - \widehat{y}_t$, we obtain the following sectoral NKPC:

$$\pi_{kt} = \kappa_k\widetilde{y}_t + \gamma_k\widetilde{y}_{R,kt} + \beta\mathbb{E}_t\pi_{k,t+1}, \tag{2.26}$$

where $\kappa_k \equiv \lambda_k(\sigma + \frac{\varphi+\alpha}{1-\alpha})$ and $\gamma_k \equiv \lambda_k(\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})$. Alternatively, the NKPC can be rewritten as:

$$\pi_{kt} = \kappa_k\widetilde{y}_t - \eta\gamma_k\widetilde{p}_{R,kt} + \beta\mathbb{E}_t\pi_{k,t+1}. \tag{2.27}$$

As is the case in standard multi-sector NK models, sectoral heterogeneities give birth to relative price (or quantity) dispersion across sectors. Consequently, a full stabilization of both inflation and output gap is no longer feasible. Moreover, while a positive aggregate output gap raises inflation in all sectors, an increase in relative price (or quantity) in one sector has a disinflationary impact in that sector and increases inflation pressure in the other sectors.

Derivation of the Welfare Loss Function

The second order Taylor expansion of the representative household's utility U_t around a steady-state (C, L) in terms of log deviations can be written as:

$$U_t - U \approx U_c C\Big(\widehat{y}_t + \frac{1-\sigma}{2}\widehat{y}_t^2\Big) + \sum_{k=1}^{K} U_{L_k} L_k \Big(\widehat{l}_{kt} + \frac{1+\varphi}{2}\widehat{l}_{kt}^2\Big) di.$$

Note that

$$(1-\alpha)\widehat{l}_{kt} = \widehat{y}_{kt} - a_{kt} + d_{kt},$$

where $d_{kt} \equiv (1-\alpha)\int log(\frac{P_{kt}(i)}{P_{kt}})^{-\frac{\varepsilon_k}{1-\alpha}} di.$

Proposition 2.7.1 $d_{kt} = \frac{\varepsilon_k}{2\Theta} var_i\{p_{kt}(i)\}$, *with* $\Theta_k \equiv \frac{1-\alpha}{1-\alpha+\alpha\varepsilon_k}$

Proof: Gali (2015) Chapter 4 Therefore,

$$U_t - U \approx U_c C\Big(\widehat{y}_t + \frac{1-\sigma}{2}\widehat{y}_t^2\Big) + \sum_{k=1}^{K} \frac{U_{L_k} L_k}{1-\alpha}\Big(\widehat{y}_{kt} + \frac{\varepsilon_k}{2\Theta_k} var_i\{p_{kt}(i)\} + \frac{1+\varphi}{2(1-\alpha)}(\widehat{y}_{kt} - a_{kt})^2\Big) + t.i.p.,$$

where $t.i.p$ denotes the terms independent of policy. Under the assumption that cost of employment is subsidized optimally at sectoral level to eliminate distortions originating from monopolistic competition, the steady-state is efficient and $-\frac{U_{L_k}}{U_c} = MPN$.

Approximate the CES aggregate C_t defined in (2.16) around $c_k = c + log(n_k)$:

$$\sum_{k=1}^{K} n_k \widehat{y}_{kt} \approx \widehat{y}_t - \frac{1-\eta^{-1}}{2}\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2,$$

with $\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2 \equiv \sum_{k=1}^{K} n_k(\widehat{y}_{kt} - \widehat{y}_t)^2$. Using the fact that $MPN =$

$(1-\alpha)(Y_k/L_k), Y = C$, it follows that:

$$
\begin{aligned}
\frac{U_t - U}{U_c C} &\approx -\frac{1}{2}\Big[\sum_{k=1}^{K}\Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\}\Big) - (1-\sigma)\widehat{y}_t^2 - (1-\eta^{-1})\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2 \\
&+ \frac{1+\varphi}{1-\alpha}\sum_{k=1}^{K} n_k(\widehat{y}_{kt} - a_{kt})^2\Big] + t.i.p \\
&= -\frac{1}{2}\Big[\sum_{k=1}^{K}\Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\}\Big) - (1-\sigma)\widehat{y}_t^2 - (1-\eta^{-1})\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2 \\
&+ \frac{1+\varphi}{1-\alpha}\sum_{k=1}^{K} n_k(\widehat{y}_{kt}^2 - 2\widehat{y}_{kt}a_{kt})\Big] + t.i.p \\
&= -\frac{1}{2}\Big[\sum_{k=1}^{K}\Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\}\Big) + (\sigma + \frac{\varphi+\alpha}{1-\alpha})\widehat{y}_t^2 \\
&+ (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2 - 2\frac{1+\varphi}{1-\alpha}\sum_{k=1}^{K} n_k \widehat{y}_{kt}a_{kt}\Big] + t.i.p \\
&= -\frac{1}{2}\Big[\sum_{k=1}^{K}\Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\}\Big) + (\sigma + \frac{\varphi+\alpha}{1-\alpha})\widehat{y}_t^2 \\
&+ (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2 - 2\frac{1+\varphi}{1-\alpha}\sum_{k=1}^{K} n_k(\widehat{y}_t + \widehat{y}_{kt} - \widehat{y}_t)a_{kt})\Big] \\
&= -\frac{1}{2}\Big[\sum_{k=1}^{K}\Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\}\Big) + (\sigma + \frac{\varphi+\alpha}{1-\alpha})\widehat{y}_t^2 \\
&+ (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})\sum_{k=1}^{K} n_k \widehat{y}_{R,kt}^2 - 2(\sigma + \frac{\varphi+\alpha}{1-\alpha})\widehat{y}_t y^N \\
&- 2(\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})\sum_{k=1}^{K} n_k(\widehat{y}_{kt} - y_t)(\widehat{y}_{kt}^N - y_t^N)\Big] + t.i.p \\
&= -\frac{1}{2}\Big[\sum_{k=1}^{K}\Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\}\Big) + (\sigma + \frac{\varphi+\alpha}{1-\alpha})\widetilde{y}_t^2 \\
&+ (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha})\sum_{k=1}^{K} n_k \widetilde{y}_{R,kt}^2\Big] + t.i.p,
\end{aligned}
$$

where $\widetilde{y}_t \equiv y_t - y_t^N$. From line 2 to line 3, we have used the fact that$\sum_{k=1}^K n_k \widehat{y}_{kt}^2 = \sum_{k=1}^K n_k \widehat{y}_{R,kt}^2 + (\sum_{k=1}^K n_k \widehat{y}_{kt})^2 \approx \sum_{k=1}^K n_k \widehat{y}_{R,kt}^2 + \widehat{y}_t^2$. From line 4 to line 5, where the fact was used that $a_{kt} = \frac{\sigma(1-\alpha)+\alpha+\varphi}{1+\varphi} y_t^N$ and $a_{kt} - \sum_{k=1}^K a_{kt} = \frac{-\eta^{-1}(1-\alpha)+\alpha+\varphi}{1+\varphi} \widehat{(} y_{kt}^N - y_t^N)$.

To summarize, the second order approximation of the representative consumer's welfare loss as a fraction of steady-state consumption is:

$$\begin{aligned} W &= \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Big(\frac{U_t - U}{U_c C} \Big) \\ &= -\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Big[\sum_{k=1}^{K} \Big(\frac{\varepsilon_k n_k}{\Theta_k} var_i\{p_{kt}(i)\} \Big) + (\sigma + \frac{\varphi+\alpha}{1-\alpha}) \widetilde{y}_t^2 \\ &+ (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha}) \sum_{k=1}^{K} n_k \widetilde{y}_{R,kt}^2 \Big] + t.i.p. \end{aligned}$$

Proposition 2.7.2 $\sum_{t=0}^{\infty} \beta^t var_i\{p_{kt}(i)\} = \frac{\theta_k}{(1-\beta\theta_k)(1-\theta_k)} \sum_{t=0}^{\infty} \beta^t \pi_{kt}^2$

Proof: Woodford, 2011 Chapter 6

Thus we obtain the following welfare loss function:

$$\begin{aligned} W = &-\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Big[\sum_{k=1}^{K} \frac{\varepsilon_k}{\lambda_k} n_k \pi_{kt}^2 + (\sigma + \frac{\varphi+\alpha}{1-\alpha}) \widetilde{y}_t^2 \\ &+ (\eta^{-1} + \frac{\varphi+\alpha}{1-\alpha}) \sum_{k=1}^{K} n_k \widetilde{y}_{R,kt}^2 \Big] + t.i.p., \end{aligned}$$

where $\lambda_k \equiv \frac{(1-\beta\theta_k)(1-\theta_k)}{\theta_k} \Theta_k$ defined as above. Normalize the weights on π_{kt} such that $\sum \omega_k = 1$:

$$W = -\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Big[\sum_{k=1}^{K} \phi_k \pi_{kt}^2 + \lambda_y \widetilde{y}_t^2 + \lambda_{R_y} \sum_{k=1}^{K} n_k \widetilde{y}_{R,kt}^2 \Big] + t.i.p., \tag{2.28}$$

where

$$\phi_k = \frac{n_k \varepsilon_k \lambda}{\lambda_k}, \qquad \lambda_y = (\sigma + \frac{\varphi + \alpha}{1-\alpha})\lambda, \qquad \lambda_{Ry} = (\eta^{-1} + \frac{\varphi + \alpha}{1-\alpha})\lambda,$$

and λ is defined as:

$$\lambda \equiv (\sum_{0}^{K} n_k \varepsilon_k \lambda_k^{-1})^{-1}.$$

From the sectoral demand equation, one can rewrite sectoral output dispersion as a function of sectoral price dispersion:

$$W = -\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Big[\sum_{k=1}^{K} \frac{\varepsilon_k}{\lambda_k} n_k \pi_{kt}^2 + (\sigma + \frac{\varphi+\alpha}{1-\alpha})\widetilde{y}_t^2 + \eta(1 + \frac{\varphi+\alpha}{1-\alpha}\eta) var_k(\widetilde{p}_{kt}) \Big] + t.i.p.$$

Normalize the weights on π_{kt}:

$$W = -\frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \Big[\sum_{k=1}^{K} \phi_k \pi_{kt}^2 + \lambda_y \widetilde{y}_t^2 + \lambda_{R_p} var_k(\widetilde{p}_{kt}) \Big] + t.i.p., \quad (2.29)$$

where

$$\lambda_{Rp} = \eta(1 + \frac{\varphi+\alpha}{1-\alpha}\eta)\lambda.$$

3

TIME-VARYING FORECAST PERFORMANCE OF CORE INFLATION

This decomposition [into core and non-core inflation] is useful because food and energy prices can be extremely volatile, with fluctuations that often depend on factors that are beyond the influence of monetary policy, such as technological or political developments (in the case of energy prices) or weather or disease (in the case of food prices). As a result, core inflation usually provides a better indicator than total inflation of where total inflation is headed in the medium term.

— Janet L. Yellen, *The Philip Gamble Memorial Lecture*
September 24, 2015

3.1 INTRODUCTION

Central banks are concerned to know how underlying inflation will evolve. Their price stability goals are often defined in terms of inflation over the medium term. For instance, the price stability objective of the ECB – as clarified by the Governing council in 2003 – is to maintain year-on-year increases in the Harmonised Index of Consumer Prices (HICP) for the euro area of below, but close to, 2% *over the medium term.*[1] Nevertheless, the volatility in headline inflation can pose a challenge for policymakers to detect the effects of cyclical inflation pressures. To succeed in this exercise, central banks need to discern persistent sources of inflationary pressures from short-lived, reversible movements. Transitory variation potentially complicates both inflation forecasting and policymaking because it is unrelated to changes in cyclical inflation pressures, or it might deteriorate the public's confidence in the central bank's commitment to long-run price stability.

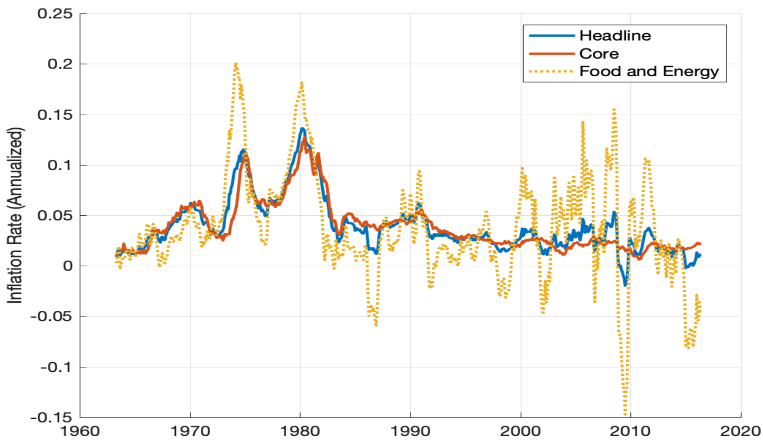
Measures of core inflation are designed to filter out the short-term volatility in headline inflation in order to reveal the cyclical signal (underlying inflation).[2] The most common of these core measures are exclusion indices, e.g., consumer price inflation excluding food and energy (CPIExFE). They are constructed by identifying several items that are considered to be the cause of the excess volatility and then building a new price index that excludes them throughout history. Historically, these tend to be food and energy prices.[3] Figure 3.1 illustrates the idea of removing volatile food and energy prices (yellow

[1]Equally, in their statement on longer-run goals and monetary policy strategy, the U.S. Federal Reserve's Federal Open Market Committee (FOMC) "reaffirms its judgment that inflation at the rate of 2 percent, as measured by the annual change in the price index for personal consumption expenditures, is most consistent over the longer run with the Federal Reserve's statutory mandate" since January 24, 2012.

[2]The concept was conceived by Gordon (1975) to describe an underlying instead of a transitory inflation rate.

[3]Blinder (1997) has alternatively argued that the real reason they were excluded was that they are mostly beyond the control of the central bank.

FIGURE 3.1: Inflation Comparison: Headline vs. Core



Note: Comparison of different measures of inflation. Headline CPI inflation (year-on-year change) in solid blue line, core CPI inflation in solid red line and Food and Energy CPI inflation in yellow dashed line.

dotted line) from headline inflation (solid blue line) in order to obtain an estimate of underlying inflation (solid red line) for CPI. Another way to construct an index of core inflation[4] are temporary exclusion indices (or central tendency statistical measures) such as trimmed-mean or weighted median.[5] They are based on the idea that large price changes in a few items can be the source of excess volatility and should therefore be excluded (Bryan and Pike, 1991 and Bryan and Cecchetti, 1994). In this paper, I focus the analysis on these two because they dominate the portfolio of monitored core inflation rates by most central banks.[6] Because of their simplicity, they are more transparent and easier to

[4]Alternative approaches to constructing core inflation include persistence weighting, variance weighting, component smoothing, exponential smoothing, and dynamic factors, among others. For an overview of different core measures, see the extensive account in Detmeister (2011).

[5]Examples used in praxis are the Dallas Fed trimmed-mean PCE and the FRB of Cleveland's median CPI.

[6]Ehrmann et al. (2018) provide an overview in Table A.

communicate to the public.

A crucial assumption behind the use of core as a measure of underlying inflation is that the excluded time series are indeed noise. If those items had a long-lasting impact on medium term inflation, removing them from inflation would disregard their signaling effect. With less information, forecasts from core inflation would then perform worse than alternative inflation measures. How well does core inflation forecast medium term inflation? Do alternative measures of core contain different information sets and therefore perform differently? Does the relative performance of those core measures change over time? What is the role of the signaling effect of non-core inflation?

This paper revisits the empirical evidence on these questions by adding an important dimension to the analysis: I allow for instabilities. There is ample evidence on the existence of parameter instability in forecasting GDP or inflation (Stock and Watson, 2003, Clark and McCracken, 2005, or Faust and Wright (2013)). I show that instabilities are also present when forecasting underlying inflation. Existing studies follow classic approaches (Diebold and Mariano, 1995 or West, 1996) and compare competing models' relative forecasting performance on average over different periods. In contrast, I follow Giacomini and Rossi (2010) and apply their Fluctuation test to check for equal predictive accuracy by considering that the relative performance of competing models might have changed over time. In practice, the Fluctuation test uses the test statistic of Diebold and Mariano (1995) computed over rolling windows to test the null hypothesis that the forecast never beats the benchmark at any point in time. In this paper, I focus on predicting medium term inflation as the three-year-ahead inflation rate[7] by using a forecasting model that includes different core inflation measures as explanatory variables and possible controls. I study different measures of inflation such as consumer price inflation (CPI) and personal consump-

[7] I check for robustness to alternative definitions such as 4-year-ahead inflation rate or year-on-year inflation in two years.

tion expenditure inflation (PCE) as well as various measures of core measures inflation as predictors, e.g., exclusion indices (e.g., PCEExFE) or statistical tendency (Dallas FED trimmed-mean).

Our empirical findings confirm that there are instabilities in forecasting medium term inflation. I document three distinct periods when comparing the forecast performance of core inflation, in comparison to headline inflation. In the 1970s and until the middle of the 1980s, headline outperforms core in predicting underlying inflation. However, the predictive ability reversed, and core inflation had a better relative forecasting performance from the mid-1980s until the beginning of the 2000s. In the later part of the sample and recent years, there is no statistical difference between both forecast models.

These results are consistent with the changes in the (time-series) process of inflation that have been documented in Mishkin (2007) or Stock and Watson (2007). Those studies estimate a steady decline in the persistence of inflation after the mid-1980s. Fluctuations in inflation tended to fade away more quickly. Since core inflation abstracts from those transitory components, its relative performance increased compared to autoregressive models. In recent years, inflation is harder to forecast in the sense that competing models perform equally well because successful inflation anchoring and forward-looking monetary policy have increased the importance of inflation expectations as a relevant predictor (Faust and Wright, 2013).

The results are the same when I consider alternative measures of core inflation. In detail, I redo the analysis with temporary exclusion indices (central tendency measures): (i) Dallas FED trimmed-mean PCE and FRB of Cleveland (ii) median CPI and (iii) 16-percent trimmed-mean CPI. The availability of those indices is restricted to a shorter sample. Notwithstanding, the evidence reports that these core measures predict better than headline inflation in the 1990s and until the beginning of the 2000s. In recent times, there is no statistical difference between forecast models.

This paper also studies which alternative core measure performs best. Previous studies (Dolmas and Koenig, 2019, or Luciani and Trezzi, 2019, or Ehrmann et al., 2018 for Europe) could not find clear evidence that any alternative core inflation measure has a higher forecast accuracy than permanent exclusion indices. According to my analysis, the reason for the ambiguity is that the relative forecast accuracy of different core measures greatly varied over time. For instance, trimmed PCE was a better predictor than PCE excluding food and energy (PCEExFE) in the 1990s, but since then, the relative forecast performance has reversed. There is no single core inflation measure that outperforms the others all the time. Instead, different measures perform well at different times. Finally, I show that averaging – looking at 36-month inflation instead of year-on-year – may not be a successful way to construct core inflation. Once I take instabilities into account, the results do not unambiguously confirm the hypothesis that longer samples provide better forecasts in contrast to previous studies (Detmeister, 2011, or Bryan and Meyer, 2011). First, the relative forecasting performance of different sample lengths highly depends on the forecasting model considered. Second, averaging does reduce not only volatility but also eliminates signals. The results suggest that the transitory component is not only noise. Instead, it may at times contain signals about future inflationary trends missed by averaging.

For this reason, I test for the importance of the signaling effect of non-core components for underlying inflation and if it has changed over time. Because core only predicts better in the middle sample, this raises the question of whether the excluded items contain essential information (signals) about future inflation. To test for the signaling effect, I apply a forecast rationality test to detect if forecasts from core inflation are not rational, e.g., biased, and information from non-core inflation could explain the lack of rationality. Given the instabilities in forecasting medium term inflation that I have uncovered before, situations with lack of rationality might appear in sub-samples of the data. Thus, I apply a rationality test that is robust to the presence of instabilities as outlined by Rossi and Sekhposyan (2016).[8] I find that non-core components contain important information about future medium term inflation that can complement forecasts from core inflation. The evidence indicates that there was a signaling channel in the 1980s and early 2000s.[9] However, the signaling effect does not explain why there is no difference in predictive accuracy between core and headline in recent years.

Related Literature. The paper is closely related to two branches of literature. More broadly, this paper adds to the literature that has focused on forecasting inflation and instabilities. Stock and Watson (2007) show that inflation has become both harder and easier to forecast over time. They argue that while the overall MSFEs of models has fallen, the difference between models became smaller. The authors relate this observation to the declining persistence of inflation in a stochastic trend-cycle model. Faust and Wright (2013) provide a horse-race between a large set of models and methods. They find that judgmental survey forecasts outperform most other models. A number of papers have documented instabilities in forecasting GDP and inflation (e.g. Stock and Watson, 2003, Clark and McCracken, 2005, or Nason, 2006), and different methods to deal with those have been developed (e.g. Giacomini and Rossi, 2010). Rossi and Sekhposyan (2016) applied these frameworks to show how the forecasting performance of different economic models for U.S. inflation and GDP growth has changed over time. They report that most indicators lost their predictive power for inflation in the 1980s. I show that instabilities also play an important role when forecasting medium term inflation. I provide empirical evidence that supports the

[8] A similar framework was used by Hoesch et al. (2020) to study the informational advantage of central banks with respect to the private sector.

[9] The first can be linked to the oil price shocks in the 1970s and the second to the run-up of commodity prices in the mid-2000s.

previous estimates of a change in forecasting performance in the 1980s and adds another reversal at the beginning of the 2000s.

Second, this paper is related to a literature that studies the forecasting performance of different measures of core inflation, e.g. Blinder and Reis (2005), Smaghi (2011), Bryan and Meyer (2011), Crone et al. (2013), Thornton (2011), Krugman (2011), Lenza and Reichlin (2011). I also compare the relative forecasting performance of core and headline, but in contrast, I focus on out-of-sample performance. Out-of-sample accuracy is considered because in-sample analyses are subject to overfitting and structural breaks. This often means that good in-sample fit fails to translate into good out-of-sample forecasting performance. Detmeister (2011), Detmeister (2012), Dolmas and Koenig (2019) or Luciani and Trezzi (2019) compare different measures of core inflation with regards to a number of different criteria including volatility, size of revisions and predictive power. I focus solely on predictive accuracy because it is the most important criterium. Moreover, some of the alternative core measures are constructed to outperform core on specific criteria (e.g., volatility), which would make a comparison unnecessary. In contrast to these studies, I consider the role of instabilities in forecasting performances. Rich and Steindel (2005) Rich and Steindel (2007) for the U.S. and Ehrmann et al. (2018) for the E.U. also identify instabilities, noting that the performance of the measures of underlying inflation in tracking the persistent component of headline inflation is episodic. I show that by focusing on the average performance over subsamples, we might average over periods of under- and overperformance and lose important information on the evolution of relative forecast performance of different models over time. Moreover, the present study considers a longer sample period and reports evidence on multiple break dates.

The paper is organized as follows. Section 3.2 discusses the motivation to look at time-variation in the forecast performance of core inflation. Section 3.3 provides an overview of the empirical framework and data. It also discusses results on the forecast performance of exclusion indices.

Section 3.4 summarizes findings for other core inflation measures, and Section 3.5 investigates if non-core inflation components add information to core forecasts. Section 3.6 provides details on robustness exercises and Section 3.7 concludes.

3.2 ARE INSTABILITIES IMPORTANT IN FORECASTING MEDIUM TERM INFLATION?

In this section, I compare the pseudo out-of-sample forecasting performance of different models using either core or headline inflation for predicting future U.S. medium term inflation. Using rolling window estimates of loss differences, I confirm that the relative forecast performance is indeed time-varying. I discuss that the forecast performance of different core measures may vary over time due to (i) a changing persistence of inflation in the excluded items or (ii) a changing composition and volatility of those goods. First, I describe the forecasting model and the data considered that will be used in this and the next section.

3.2.1 *Forecasting Model*

In order to investigate whether core or headline inflation is the better predictor of future medium term inflation, I compare the forecasting performance of regressions of the form

$$\pi_{t,t+h} = \alpha + \theta(L)x_{t-12,t} + \varepsilon_t,\ t = 1, 2, ...T, \tag{3.1}$$

where $\pi_{t,t+h}$ is a measure of medium term inflation, $x_{t-12,t}$ is the year-on-year change in a predictive variable and α is a constant. Specifically, h is the forecasting horizon. The model to predict inflation differs slightly from other papers in this literature which either use a form of predicting medium term inflation via $\pi_{t,t+h} = \alpha + \theta x_{t-12,t} + \varepsilon_t$ (Blinder and Reis, 2005 and Crone et al., 2013) or $\pi_{t,t+h} - \pi_{t,t-s} = \alpha + \theta(x_{t,t-12} - \pi_{t,t-s}) + \varepsilon_t$ as in Clark (2001), Cogley (2002), Detmeister (2011), Detmeister (2012) or Rich and Steindel (2007). I specify the model in the above form to ackowledge results from Stock and Watson (2007) or Faust and Wright (2013) that stress the importance of autoregressive terms. Faust and Wright (2013) show that it is hard to beat a univariate AR(1) model for forecasting inflation. Only surveys (Greenbook, SPF, Blue Chip) or forecast combinations (Rossi (2013)) consistently beat an AR(1) in gap form with a fixed slope coefficient. Notwithstanding, I will also consider these alternative forecast models in the robustness section of this paper.

Moreover, Rossi and Sekhposyan (2010) detail that there are few explanatory variables for inflation that still have significant forecasting power for inflation after the 1980s, justifying the exemption of those in the baseline analysis. Nevertheless, to acknowledge the literature on Phillips curves, I also consider models including additional explanatory variables related to economic activity. I add the output gap (CBO) as an additional control In one model as in standard Phillips curve regressions.

The measure of medium term inflation I consider as baseline is the (annualized) headline inflation rate over the next 36 months, i.e., $h = 36$ or $\pi_{t+36,t} = 100/3 \ln(CPI_t/CPI_{t-12})$. Nevertheless, I also consider inflation over the next 48 months or year-on-year inflation from 24 months ahead to 36 months ahead. As a possible explanatory variable, $x_t, t-12$ is either the year-on-year change in (i) headline inflation or (ii) core inflation. As a measure of headline inflation, I consider year-on-year (i) consumer price inflation (CPI), (ii) personal consumption expenditure inflation (PCE), and (iii) producer price inflation (PPI). Inflation is constructed as annualized change in prices, for instance, for the y-to-y inflation rate as $\pi_{t,t-12} = 100 \ln(CPI_t/CPI_{t-12})$. In contrast, I consider the corresponding exclusion inflation index as measure of core inflation, i.e., the respective change in prices from the index excluding food and energy prices – (i) CPIExFE (ii) PCEExFE and (iii) PPIExFE. In the robustness section 3.6, I also study the performance of food and energy inflation and add controls to equation (3.1).

Model (3.1) is estimated by OLS in rolling samples of 120 observations (10 years). Thus, as the sample starts in 1959:1, and the effective sample size reduces by one observation due to differencing, the first 36-month ahead out-of-sample forecast is made for 1972:2.

Next, to compare relative forecast performance, I construct relative Mean Squared Forecast Errors (rMSFE). Therefore, I first denote the pseudo out-of-sample forecast errors of model (3.1) with core or headline inflation, respectively, by $\widehat{\varepsilon}_{t+h}^{core}$ and $\widehat{\varepsilon}_{t+h}^{headline}$. They are constructed as the difference between the realization of medium term inflation $\pi_{t,t+h}$ and its forecasted value using the model in (3.1) with either core or headline as predictor $\widehat{\pi}_{t+h|core}$ or $\widehat{\pi}_{t+h|headline}$. Next, I construct estimates of the relative Mean Squared Forecast Errors (rMSFE) as the difference between the mean squared forecast errors of both models. In order to capture time variation in relative performance, I construct those rMSFEs in a rolling window fashion:

$$rMSFE_t = \left(\frac{1}{m} \sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{headline})^2 - \frac{1}{m} \sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{core})^2 \right), \qquad (3.2)$$

where m is the size of those windows. The construction of the loss differences depends on the choice of two parameters: (i) the forecast evaluation window size (m), and (ii) the size of the sample (R) used to estimate the forecasting model. There is a trade-off between good estimates for the forecasting model and accurate estimates of rMSFE vis-á-vis a large sample of rMSFE estimates to better track the evolution of relative forecast performance over time. I chose the estimation sample size of 120 months and window size, m, of 120 months in the benchmark case. Nevertheless, I verified the robustness of those results to alternative choices for both parameters in Section 3.6.

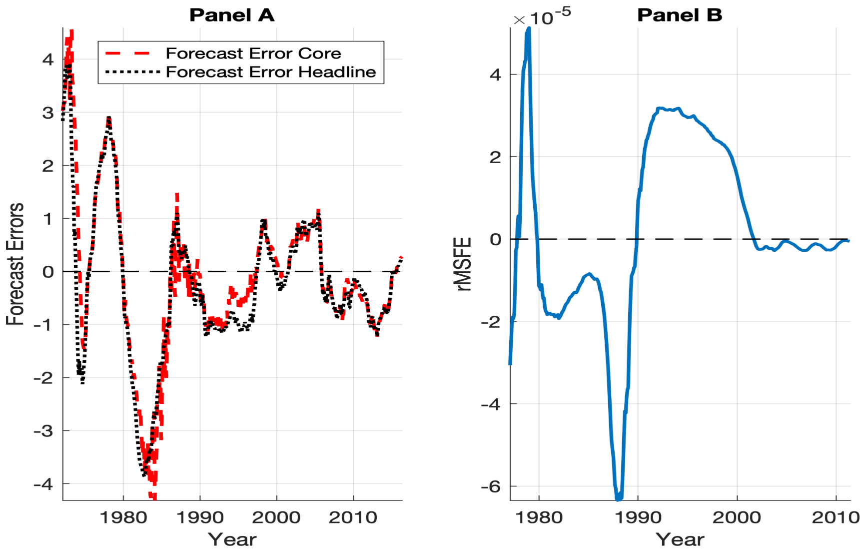
3.2.2 *Instabilities in Forecasting Medium Term Inflation*

An important challenge that an econometrician who attempts to investigate which measure of core inflation predicts better faces, is that forecasts are unstable. For instance, there is broad empirical evidence on the existence of parameter instability in forecasting GDP and inflation (as documented, for example, by Stock and Watson, 2003, and Clark and McCracken, 2005 or Nason, 2006, or on how economic models' forecasting performance for U.S. output growth and inflation changed over time (Rossi and Sekhposyan, 2016) but also when using measures of core inflation (Detmeister, 2011, or Ehrmann et al., 2018). The presence of instabilities in forecasting medium term inflation is documented in Figure 3.2. Panel A of Figure 3.2 reports forecast errors from predicting medium term inflation as the difference between the realization of medium term inflation and its predicted value. The dashed red line reports forecast errors associated with core inflation, while the black dotted line shows the forecast errors using headline inflation as a predictor.

The first observation from Figure 3.2 is that medium term inflation was harder to forecast at the beginning of the sample, as indicated by the larger forecast errors. This was equally true for both predictors. In recent years, however, the forecast errors and the difference in forecast errors have become smaller. This confirms the findings in Stock and Watson (2007) that inflation became both easier and harder to forecast over time and extends them to medium term inflation.

From the forecast errors, it becomes evident that there are times when the forecasts from both explanatory variables consistently over- or underpredict medium term inflation. For instance, in the 1990s, both headline and core inflation systematically overpredicted medium term inflation, i.e., the forecasts were likely biased, and the forecast models misspecified. Nevertheless, even during these times, one model can still perform better than the other. From the graph, it is evident that the forecast errors from core inflation in the 1990s were smaller. This observation is confirmed when we look at the rMSFEs in Panel B of 3.2. Intuitively, the graph shows a smoothed version of the difference in forecast errors. Positive values reflect that core inflation has lower

FIGURE 3.2: Instability in Forecasting Medium Term Inflation



Note: This figure compares forecasts of medium term inflation from using core inflation or headline inflation as predictors. The left Panel A shows forecast errors as actual in predicted value of medium term inflation over time. Forecast errors from core inflation are depicted with a red dashed line and those from headline inflation with a black dashed line. The right Panel B reports rolling window estimates of forecast losses as defined in Equation (3.2).

forecast errors. Again, there is evidence of time-variation in the relative forecast performance.

The suggestive evidence raises concerns for the use of full sample analyses. If the historical period includes episodes of relative over- and under-performance of a model, then on average, it might look as if the model does not have an advantage compared to the other model. However, in reality, there is a forecasting edge of that model for certain periods. Thus, I investigate the properties of the forecasts by employing statistical tests that account for the identified instability in relative forecast performance (Giacomini and Rossi, 2010). In the next section, I will describe the statistical framework used in detail. However, before, I want to briefly discuss potential explanations for why the relative forecast performance of core and headline inflation can vary over time.

3.2.3 *Why Did the Forecast Performance of Core Inflation Vary Over Time?*

According to the idea of core inflation, it is a better measure of underlying inflation because it strips out transitory variations in inflation and better reveals the cyclical state from inflation. The first underlying premise is that the excluded components have no long-lasting impact on underlying inflation. Otherwise, changes in prices of excluded items could affect medium term inflation, and including this information would result in better forecasts. Another critical assumption is that the "correct" items have been removed. The most volatile items are not necessarily the ones with the most noise, and the volatility of items can change over time. Evidence points towards time-variation in the validity of both of these two assumptions.

Excluded items can have a lasting impact on overall inflation. Historically, oil shocks were viewed to cause only transitory changes ("blip") to the inflation process (Blinder and Reis, 2005). They do not affect long-term inflationary expectations and fade away on their own and, therefore, should be ignored by central banks. However, excluded items can have a lasting impact on overall inflation if there is a pass-through from shocks to those components to headline inflation. This pass-through could work via indirect effects (i) by affecting other items in the consumer price basket through higher production costs (the network channel), (ii) by influencing inflation expectations (Coibion and Gorodnichenko, 2015) or (iii) by impacting wages. Empirical evidence confirms time-variation in this pass-through. There is widespread evidence that the relationship has been unstable over time (see, e.g., Edelstein and Kilian, 2009, Herrera and Pesavento, 2009, Blanchard and Galí, 2009, Ramey and Vine, 2010, Baumeister and Peersman, 2013). In particular, several researchers have noted a substantial decline in the macroeconomic consequences of oil price shocks. If the persistence of excluded items has changed, as this evidence suggests, it may distort the reliability of signals of future underlying inflation from core inflation.

On the other hand, core inflation may exclude the "wrong" items. Excluding the most volatile items from headline inflation might reduce not only noise but also signals. This is equally true for exclusion indices as for temporary exclusion indices like statistical central tendency measures. One channel is indirect effects, as argued before. Another could be that excluded items contain leading information (see Giannone et al., 2014). More volatile prices – and those with large price changes – tend to be more "sticky" items. As these prices change less often, the price-setters are more forward-looking and, thus, tend to incorporate persistent changes rather than transitory fluctuations.[10] But, not all volatile items are sticky prices. Moreover, the empirical evidence points towards hybrid inflation dynamics, i.e., forward- and backward-looking. In this case, it might even take longer for persistent shocks to be reflected in "sticky" prices. Instead, flexible prices could be leading indicators and improve forecast performances.

Furthermore, the volatility and expenditure share of items can change over time. Food and energy inflation indeed tend to be more volatile than core inflation, as illustrated by Figure 3.1. However, Figure 3.1 also shows that the volatility of different items can vary over time. Food and energy inflation appears to deviate less from headline and core inflation in the 1980s and 1990s. However, as outlined by Dolmas and Koenig (2019), not all sub-categories among the food and energy prices are the most volatile items. Their study reports that about 14 percent of food and energy items (by expenditure share) are less volatile than other core items. This means that we might involuntarily remove signals in the construction of core inflation, and likewise allow some excess volatility

[10]This reasoning motivates Aoki (2001)'s optimal price index that puts larger weight on more sticky prices. The FRB of Atlanta publishes such a sticky-price CPI index.

to remain in the index. This concern becomes even more alarming when we contemplate that (i) the expenditure shares and (ii) volatilities of those items are time-varying. Thus, what would have been an optimal exclusion in the past may no longer be today.

The forecast performance of different core measures may vary over time due to a changing persistence of inflation in the excluded items or a changing composition and volatility of those goods. This is why, in contrast to most of the current literature, I use a testing approach that is fully robust to instabilities. This is important as we might otherwise fail to detect a break in forecast performance.

3.3 Does Core Forecast Medium Term Inflation Better?

In this section, I compare the pseudo out-of-sample forecasting performance of different models using either core or headline inflation for predicting future U.S. medium term inflation. I explicitly account for time-variation and possible instabilities in relative forecast accuracies by performing fluctuation tests (Giacomini and Rossi, 2010). Before I turn to the results of this exercise that confirm that the relative forecast performance is indeed time-varying, I describe the forecasting model, the data considered, and the statistical method used to study the research question.

3.3.1 *Tests for Instability in Relative Forecasting Performance*

This section describes the statistical method used to study more structurally the time-variation in relative forecasting performance and its implementation. To test whether the two models' relative forecasting performance has changed over time, I employ as statistical method the

Fluctuation test developed by Giacomini and Rossi (2010).[11]

The test statistic relies on a normalized version of the rolling MSFE introduced in the last section. In particular, the Fluctuation test statistic is a measure of the local relative forecasting performance of two models over rolling windows of data and re-scaled by an asymptotic variance

$$F_{t,m}^{OOS} = \widehat{\sigma}^{-1} m^{-1/2} \left(\sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{headline})^2 - \sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{core})^2 \right), \quad (3.3)$$

where $\widehat{\sigma}^{-1}$ is a Heteroskedasticity and Autocorrelation Consistent (HAC) estimator of the asymptotic variance.[12] I consider the following bandwith estimator (Newey and West, 1987) for the asymptotic variance

$$\widehat{\sigma}^2 = \sum_{i=-q(P)+1}^{q(P)-1} \frac{(1 - |\frac{i}{q(P)}|)}{P} \sum_{j=R+h}^{T} ((\widehat{\varepsilon}_j^{head})^2 - (\widehat{\varepsilon}_j^{core})^2)((\widehat{\varepsilon}_{j-i}^{head})^2 - (\widehat{\varepsilon}_{j-i}^{core})^2), \quad (3.4)$$

where the bandwidth $q(P)$ is $q(P) = P^{1/4}$.

The null hypothesis of the test is that the two models' forecasting performance is equal at each point in time, i.e.

$$H_0 : E((\widehat{\varepsilon}_t^{headline})^2 - (\widehat{\varepsilon}_j^{core})^2)) = 0, t = R + h, ..., T. \quad (3.5)$$

Giacomini and Rossi (2010) show that the asymptotic distribution of the Fluctuation test under the H_0 can be approximated by functional Brownian motion. Moreover, they also provide critical values for various

[11] They developed a second test; the One-time Reversal test. This test can be employed to test for an exact break date. In contrast, the Fluctuation test uses local losses and suggests a time period for the change in relative performance. On the other hand, the former test is less applicable in the context of multiple break dates.

[12] The asymptotic variance is $\sigma = var(P^{-1/2} \sum_{j=R+h}^{T} ((\widehat{\varepsilon}_j^{headline})^2 - (\widehat{\varepsilon}_j^{core})^2))$ where P is the sample used for forecasting, i.e., $P = T - R$.

significance levels as well as window and sample sizes.[13] Following Rossi and Sekhposyan (2010), I chose the test statistic to represent the mid-point of the forecast evaluation window, which implies that the effective sample size reduces by $m/2$ observations at the beginning and end of the sample. Thus, the sample path for the test is $t = R + h + m/2, ..., T - m/2 + 1$ where R is the size of the estimation sample.

The test is implemented by plotting the evolution of local losses together with critical values. If local losses are outside the bands indicated by the critical values, this can be interpreted as one model outperforming the other. In our specification, positive relative local losses indicate that core performs better. Thus, if the path of the Fluctuation test crosses the upper bound, this reads that core outperforms headline at some point around this time. If the path crosses the lower bound, then one can conclude that the headline model forecasts better. Note that due to the nature of the test displaying normalized averages of the relative performance over a window m, it is not possible to infer the exact timing of a change in forecasting performance but only an indication – which is more or less precise given the choice of m (hitherto, in our case ten years).[14] Finally, it is worth noting that the applicability of the Fluctuation test relies on i) stationarity assumptions (no high persistence in the local losses, e.g., unit roots) and ii) global covariance stationarity (no breaks in the variance of MSFE).

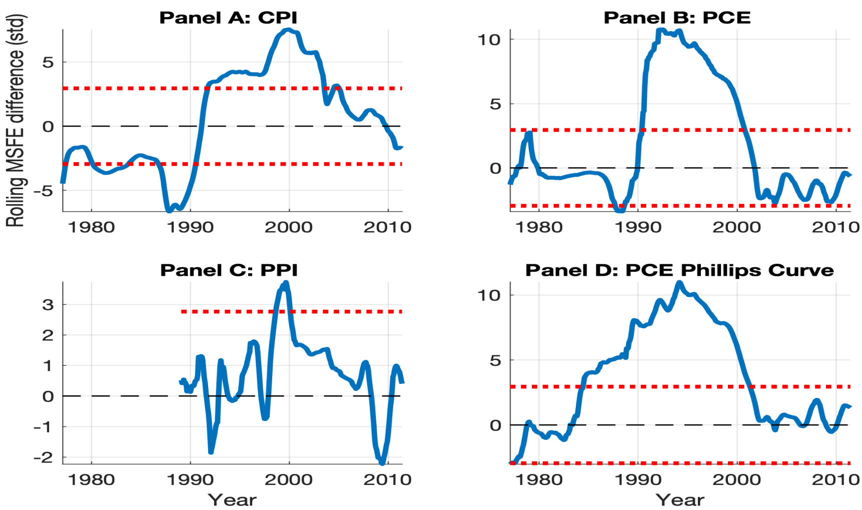
[13]I implemented the test using Giacomini and White (2006)'s framework where the losses depend on estimated in-sample parameters and do not need a correction for parameter error. The latter's requirement is the use of a rolling estimation with a fixed window size to produce the out-of-sample forecasts which I will implement.

[14]If the interest is to identify the exact timing of a break date, it is advisable to use the one-time reversal of Giacomini and Rossi (2010) which, however, is not applicable in the present case since the later evidence will indicate multiple instead of just one break date.

3.3.2 *Core Predicts Better Only in the Middle Part of the Sample*

In this section, I report results on the Fluctuation test to study the relative predictive ability of core and headline inflation to forecast U.S. medium term inflation by considering different measures of medium term inflation and the inflation rate itself. As measures of core inflation, I consider exclusion indices, i.e., inflation indices that exclude food and energy prices. As the baseline case, I will consider inflation over the next 36 months – i.e. $h = 36$ in equation (3.1) – as measure of medium term inflation.

FIGURE 3.3: Fluctuation Test: Medium Term Inflation



Note: Fluctuation Test with 36m medium term inflation. Fluctuation test statistic, calculated as (standardized) difference between MSFE of the headline inflation and MSFE of the core model calculated over rolling windows ($m = 100$), across different specifications for core inflation and equation (3.1). Red dashed line shows the fluctuation test's one sided critical value at 10%.

Figure 3.3 reports empirical evidence based on tests of equal predictive ability on average over the sample period, starting in 1977:2 and ending in 2011:6 – except for PPI inflation for which the available sample is shorter[15] and thus the rolling normalized rMSFEs start in 1989:2. For each panel, the graph shows the Fluctuation test statistic (solid blue line), which is the normalized rMSFE differences over time, together with the two-sided critical values at 10% (red dashed lines). Positive values of the test statistic indicate that the model with core inflation has higher forecast accuracy than the model with headline inflation.

Panel A focuses on consumer price inflation (CPI). By inspection of the path of the test statistic, there is compelling empirical evidence that there is strong time variation in the relative performance of the two economic models. This is consistent with the accounts of Stock and Watson (2003), Faust and Wright (2013), Rossi and Sekhposyan (2010),[16] and Detmeister (2011) that there are instabilities in the relative forecasting performance of different forecast models. In fact, the visual evidence points towards two reversals in the evolution of the two models' relative forecast performance. Next, I want to test if these differences are statistically significant. Therefore, I test the null hypothesis that the two models' relative performance is the same at each point in time. The alternative is that either model performs better at the 10 percent significance level. The red dashed line depicts the critical values for testing the null hypothesis at the 10% level. How can we interpret the results of the Fluctuation test from the graph? If the normalized rMSFE differences remain within the bounds, we cannot reject that the relative forecast performance is the same at all times, i.e., core never beats headline. Instead, if the path is outside the bounds, the relative predictive ability did not stay the same over time.

Panel A suggests that there have been broadly *three periods over time*. At the beginning of the sample and until the end of the 1980s,

[15] While headline PPI is available from 1947:4, BLS only started reporting an exclusion index for PPI from 1974:1.

[16] Rossi and Sekhposyan (2010) document that some variables had significant forecast performance in the early 1980s which subsequently disappeared.

core performs *worse* than headline inflation. This can be seen from the negative normalized rMSFE. Moreover, in at least three instances, the test statistic is below the negative boundary line. In the second period, which broadly goes from the end of the 1980s until the beginning of the 2000s, the normalized rMSFE turns positive and *exceeds* the upper boundary line, indicating that the relative forecasting performance of the two economic models has reversed. Eventually, since the beginning of the 2000s, the rMSFE does not exceed the critical values. Thus, the performance of core CPI is not statistically significantly different from that of headline. In summary, the empirical evidence suggests that core CPI inflation only performed better in the 1990s. Before, headline CPI had a higher forecast accuracy, and since the beginning of the 2000s, the relative predictive ability was about the same. In the interpretation of these results, it is important to note that the Fluctuation test is a supremum-type test, i.e., it either rejects or does not reject the hypothesis of equal forecast performance at all times. Here, the evidence shows that the test rejects the null hypothesis, which means that the two models do not forecast equally well all the time. In the interpretation of the three periods, I consider that the test statistic's signs of revert over time and get close to zero in recent years.

How do these results compare to using other measures of inflation? To answer this question, Figure 3.3 also plots the result of the Fluctuation test for other measures of core inflation. Thus, I will look at personal consumption expenditure inflation (PCE) next, depicted in Panel B. The evolution of the two PCE inflation models' relative predictive ability is very similar to the one of CPI inflation, except for two observations. In the first period, the evidence of a superior relative predictive ability of headline is less evident. The path of the (normalized) rMSFE is only once outside the negative boundary line and even positive for a short time at the end of the 1970s. The second observation is that core PCE has maintained its predictive ability for a shorter period than core CPI inflation at the beginning of the 2000s.

Panels C of Figure 3.3 illustrates the relative forecast performances of core vis-á-vis headline producer price inflation (PPI). The Fluctuation test's evolution strongly differs from the others insofar that the relative performance of core PPI becomes statistically different from headline PPI only at the end of the 1990s.

Finally, I consider a variation of the model in (3.1) by accounting for additional explanatory variables according to the Phillips curve relationship. In favor of such models, Stock and Watson (1999) and Stock and Watson (2013) found empirical evidence in favor of the Phillips curve as a forecasting tool by demonstrating that inflation forecasts produced by the Phillips curve are generally more accurate than forecasts based on other economic variables. Thus, I add the unemployment gap – the deviation of the unemployment rate from its natural rate estimate from the CBO – to the model in (3.1) for both core and headline inflation. Panel D visualizes the results of a Fluctuation test. The evolution of rMSFE is very similar to those from CPI and PCE inflation. According to this panel, core inflation appears to have maintained its predictive ability since the mid-1980s.

In summary, there is ample evidence that different exclusion indices as a measure of inflation have higher predictive accuracy than headline inflation from the end of the 1980s until the beginning of the 2000s. Outside this time period, there is a reversal in the relative forecast performance. While headline appears to perform better in the 1970s and part of the 1980s, there is no statistical difference between both forecast models after the 2000s. In the next section, I will discuss in more detail the results of these observations.

3.3.3 *Why Did the Forecast Accuracy Vary in the Three Episodes?*

The evidence of the section 3.3.2 indicates that there are three subperiods with distinctively different relative forecast performances of headline and core inflation. In this section, I will provide a brief discussion on the potential economic drivers of these results.

Figure 3.3 suggests that headline inflation was a better predictor of medium term inflation than core inflation in the 1970s and for most of the 1980s. Evidence by Mishkin (2007) or Stock and Watson (2007) shows that the inflation process was highly persistent during the 1970s and until the mid-1980s. Under those conditions, an increase in inflation caused the trend component to rise in tandem, and both stayed up. In these cases, inflation is well approximated by a low-order autoregression.

In fact, this was a time with a high inflation level, mainly driven by large increases in the price of oil triggered by the Yom Kippur war in 1973 and the Iranian revolution of 1979, respectively. In combination with more structural economic conditions such as the monetary policy stance, this resulted in an un-anchoring of trend inflation. Another potential contributing factor was that the share of oil in consumer expenditures was larger during this time. In the context of forecasting, non-core components contained signals about future inflation, and core inflation failed to pick those up, which lead to the inferior relative forecasting performance of core in the first part of the sample.

This changed in the second part of the sample, where I reported evidence of a reversal in relative forecast accuracy. From Figure 3.3 it becomes apparent that the second time period – from the mid-1980s until the beginning of the 2000s – is characterized by a falling trend in inflation and smaller fluctuations around that trend. This is confirmed by an analysis of Stock and Watson (2007), who estimate a time-varying trend-cycle model of inflation over time. While the variance in the transitory component of this model has remained constant, the variability of the permanent component has sharply decreased since the mid-1980. A change in inflation then reflected a change in the transitory component and not in the trend as in the 1970s. In consequence, fluctuations in inflation tended to fade away more quickly. Hence, inflation persistence was much lower, which explains the relatively poor performance of autoregressive models, like model (3.1) that uses lags of headline inflation. Instead, core inflation predicts better in this time period as it abstracts from the transitory components and better picks up the declining permanent component.

There are several possible reasons for the changing properties of the inflation process that reflect more fundamental changes in the economy and that are mostly captured in the literature on the great moderation (e.g., Blanchard and Simon, 2001, or Stock and Watson, 2003). The most prominent explanation is changes to the conduct of monetary policy. Central banks adopted inflation targeting strategies and, hence, committed to maintaining an explicit inflation target. This has possibly led to an anchoring of inflation expectations (Mishkin, 2007 or Bernanke, 2007) that eventually brought inflation to the target level and led to smaller reactions of both inflation and output to temporary shocks due to improvements in the policy trade-off. Other possibilities include changes to the nature of the structural shocks hitting the economy,[17] or changes in the structure of the real economy (e.g., structural transformation).

Since the 2000s and in recent years, we cannot reject both models' equal forecast performance. Looking at the rMSFE in Figure 3.2, one can observe that they are systematically close to zero. At the same time, overall forecast errors became smaller. In the words of Stock and Watson, 2003, inflation became both easier and harder to predict. This section's analysis confirms that it became harder because, since the beginning of the 2000s, it is more difficult to differentiate between forecast models. One explanation for this observation may be that inflation anchoring was completed at the end-1990s. The estimated persistent components in Mishkin, 2007 or Stock and Watson, 2007 had been decreasing until the end of the 1990s and remained mostly stable since then. The anchoring of inflation expectations led to a stabilization of trend inflation and hence a decline in inflation persistence. With inflation expectations anchored, a given shock to inflation – demand, commodity, exchange-rate – will

[17]This might be reflected by the smaller volatility of food and energy inflation in this period in Figure 3.3.

have a smaller effect on expected inflation and, thus, on trend inflation. Under such conditions, an inflationary shock – like a substantial rise in energy prices – is less likely to spill into expected inflation and, hence, trend inflation.

The role of inflation expectations has become even more important since central banks have started to use more forward guidance explicitly – through formal committee statements – or implicitly — through speeches and testimony by its members. Campbell et al. (2012) and Campbell et al. (2016) show that this was done long before the financial crisis. Therefore, in such an environment, a better measure of underlying inflation is inflation expectations. This hypothesis is confirmed by Faust and Wright (2013), who perform a horse-race among a large set of traditional and more recent forecasting methods, and find that judgemental survey forecasts – e.g., Blue Chip survey, survey of professional forecasters, or FED staff's Greenbook forecasts – perform best.

Another set of explanations for the behavior of forecasts in recent years is centered around the Lucas-critique. If the central bank uses a core measure to forecast inflation and successfully controls inflation, the core measure should lose its predictive ability. In this case, the best predictor of inflation will again be the target of the central bank (Detmeister, 2011, Rowe, 2011). As Blinder and Reis (2005) and Bernanke (2010) report, there was indeed a shift towards core inflation. At the beginning of the 2000s, the FED replaced CPI, first, with headline PCE, and, then, added core PCE. However, the assumptions surrounding this critique are strong and, thus, it is usually ignored by the core inflation literature.

A final observation is that headline and core might have become more similar in recent years. While the previous explanations rested on the premise that there is less to predict in medium term inflation in recent years, another possibility is that the differences in the information of both inflation rates have declined. Looking at the composition of PCE inflation, it becomes evident that the share of non-core components in the aggregate index has declined. While the PCE share of core components in the 1960s was 74.78%, it has increased to 86.96% in the 2010s. However, the evidence in Figure 3.3 suggests that despite this decline, headline and core displayed different patterns with extended divergences between both rates, mainly due to an increase in the volatility of non-core components.

3.4 Forecasting Performance of Alternative Measures of Core Inflation

In the first part of this paper, I have focused on one class of core inflation measures: exclusion indices. Exclusion indices attribute volatility in headline inflation to a subset of sectors (food and energy) and permanently exclude them. However, there are many ways to measure underlying inflation, and in this section, I will look at two sets of promising alternative measures of core inflation. The first set of core measures is temporary exclusion (central tendency statistical measures) indices, which by means of statistical methods, choose which of the items are to be excluded on a monthly basis. The second class includes frequency-based approaches that filter out the transitory component using averaging.

For each of these alternative measures, I will first study the relative forecasting performance of the alternative measures to headline inflation. Afterwards, I will follow more recent policy papers and compare the relative forecasting performance of exclusion indices in comparison to the alternative measures of core inflation.

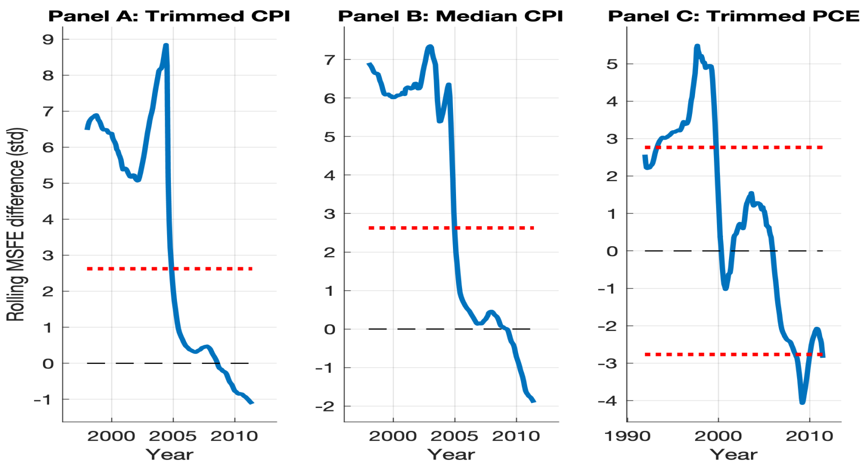
3.4.1 *Temporary Exclusion Indices*

The earliest measures of underlying inflation were exclusion indices. They build on the idea to identify those items which are the source of excess volatility in headline. They are then constructed by creating a new price index that excludes those items. The works of Bryan and Pike (1991), Bryan and Cecchetti (1994), and Dolmas (2005) have shifted the focus to the idea that extensive price changes in a few items can cause the excess volatility in headline and that those items might not be the very same throughout history. Following this idea, a core measure is then constructed by removing the items with the smallest and largest price changes each month. Notable examples for the U.S. include two measures published by the FRB of Cleveland (i) 16-percent trimmed-mean CPI inflation rate (Trimmed CPI, henceforth) and (ii) Median CPI (Median CPI, henceforth) and one from the Dallas FED (iii) trimmed-mean PCE inflation rate (Trimmed PCE). The trimmed-mean excludes fixed proportions of mass from the lower and upper tails of the distribution of item-level price changes in each period. In detail, the trimmed CPI uses symmetric trimming proportions of 8% from the lower and 8% from the upper tail. The median CPI picks the item whose expenditure weight is in the 50th percentile of the price change distribution. In contrast, the trimmed PCE uses asymmetric trimming proportions; in its calculation 24%, of the mass from the lower tail and 31 % from the upper tail is trimmed.

A number of papers (Dolmas and Koenig, 2019 or Luciani and Trezzi, 2019) have compared the relative performance of these core measures to exclusion indices. They focused on a series of characteristics besides their predictive ability for underlying inflation, including reducing volatility or the size of revisions. The consensus in these studies is that neither measure clearly dominates classical exclusion indices in a forecasting sense. In the following analysis, I will show that this result arises from the time-varying relative forecast performance of those measures.

Alternative Measures Perform Similarly. In the first exercise, I want to compare these alternative measures' forecast accuracy to headline inflation. For this purpose, I will redo the exercise of Section 3.3, where the only difference is that now I will use one of the three previous core inflation measures (i)-(iii) as a predictive variable instead of an exclusion index. Then ε^{core}_{t+h} will represent the forecast error from these models. Thus, positive values of the local losses of equation (3.3) will indicate that the respective alternative core measure performs better than headline inflation in forecasting medium term inflation. Another difference is that the sample period reduces since the alternative core measures are available for shorter timespans.[18] Thus, the respective time paths of rMSFE are 1998:1 - 2011:6 for the two CPI measures and 1991:12 - 2011:6 for trimmed PCE inflation.

Figure 3.4: Fluctuation Test: Alternative Core Inflation Measures



Note: Fluctuation Test of other core inflation measures vis-á-vis headline inflation with 36 months medium term inflation. Fluctuation test statistic, calculated as (standardized) difference between MSFE of the headline inflation and MSFE of the core model calculated over rolling windows ($m = 120$). Red dashed line shows the Fluctuation test's two sided critical value at 10%.

Figure 3.4 reports the result of the fluctuation test for the three alternative core inflation measures. The first two panels — A and B – display the results for trimmed and median CPI, and Panel C for trimmed PCE. The shorter timespan prevents an evaluation of

[18]E.g., the specific disaggregation used for trimmed PCE is only available from 1977:1, and, hence, trimmed PCE starts in this year.

predictive accuracy in the 1970s and 1980s, where according to the last section, headline inflation is expected to perform better. Focusing on the second sample period (the 1990s and early 2000s), Figure 3.4 shows that alternative core measures perform better than headline inflation. The rMSFE is positive, indicating that forecast errors using headline inflation as a predictor are larger than forecast models with core inflation. Additionally, the test statistic exceeds the upper bound, suggesting that the null hypothesis of equal performance of the two respective models can be rejected. Similarly as before, the reversal in forecast accuracy appears to be earlier for PCE inflation (beginning of the 2000s) than for CPI inflation (mid-2000s). The likely determinant of this difference is the different treatment of house prices. Overall, the results of this episode are in line with those from exclusion indices.

While the results for the alternative core CPI inflation measures in the third subsample look like those for the CPI exclusion index (CPIExFE), trimmed PCE becomes less accurate than headline inflation. Around the mid-2000s, there is a reversal in forecast accuracy of trimmed and median CPI. The (standardized) rMSFEs both turn negative. Statistically, however, we cannot reject that the alternative measures' forecast performance is different to headline CPI inflation, given that the (standardized) local losses are within the two bounds. In contrast, trimmed PCE inflation shows a briefly interrupted reversal at the beginning of the 2000s. Moreover, trimmed PCE forecasts medium term inflation worse than headline inflation at the end of the 2000s as indicated by the Fluctuation test statistic, which not only turns negative but also breaks through the lower bound. The reversal occurred around the Great Recession and extended up until the recent time period. This is in contrast to exclusion core inflation in Panel B of Figure 3.3, which never performs significantly worse than headline inflation. Does this imply that exclusion core is a better predictor of medium term inflation than trimmed PCE in recent years?

Predictive Accuracy of Various Measures of Core Inflation Varies Strongly Over Time. To answer which measure of core inflation is a better predictor of medium term inflation, I will compare the forecast performances of the alternative measures of core inflation to core exclusion indices. This means that I compare each of the previous three core measures to the respective exclusion index instead of headline inflation. Note that this is a different exercise with potentially different results than merely comparing the Figures 3.3 and 3.4. Forecast models using different core measures might have the same accuracy as headline inflation, but there could still be one model that performs better than the other.

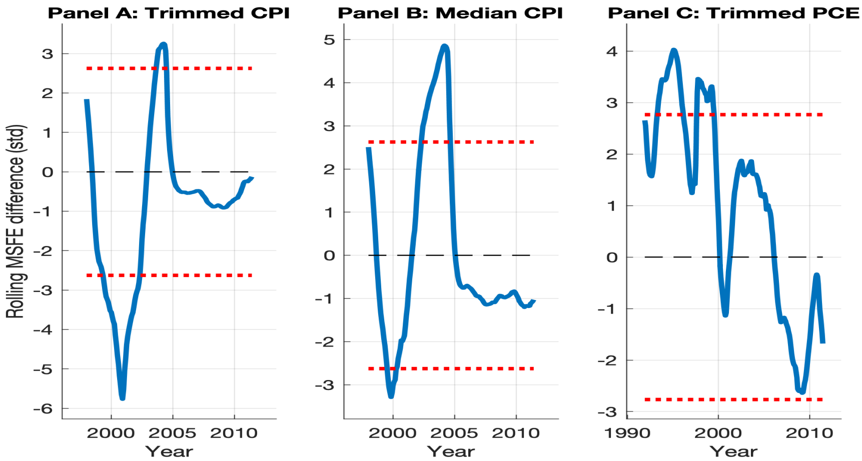
Why should we expect differences between the two core indices? As discussed in the last subsection, there are essential differences in the construction of exclusion and temporary exclusion indices. While the basket of goods removed from exclusion indices is fixed, the particular items eliminated from temporary exclusion indices are time-varying. Excluding the items with the largest price changes is likely to decrease volatility, but it might also eliminate the signals about future inflationary pressures these items contain. Massive changes in prices can still affect medium term inflation, either if they are not temporary or if they affect other items indirectly.

To test the relative predictive accuracy of trimmed and median inflation to inflation, excluding food and energy, I modify the Fluctuation test statistic from the previous section

$$F_{t,m}^{OOS} = \widehat{\sigma}^{-1} m^{-1/2} \left(\sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{excore})^2 - \sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{othercore})^2 \right), \quad (3.6)$$

where $\widehat{\varepsilon}_{t+h}^{excore}$ and $\widehat{\varepsilon}_{t+h}^{othercore}$ denote the pseudo out-of-sample forecast errors of model (3.1) with exclusion core indices or temporary exclusion indices, respectively. They are again constructed as the difference between the realization of medium term inflation $\pi_{t,t+h}$ and its forecasted value using the model in (3.1) with either trimmed or median inflation ($\widehat{\pi}_{t+h|othercore}$) or inflation excluding food and energy ($\widehat{\pi}_{t+h|excore}$) as a predictor.

FIGURE 3.5: Fluctuation Test: Alternative Core vs. Exclusion Indices



Note: Fluctuation Test alternative core inflation measures vis-á-vis exclusion indices. Fluctuation test statistic, calculated as (standardized) difference between MSFE of the excusion core inflation and MSFE of the alternative core models calculated over rolling windows ($m = 120$), across different specifications for core inflation and equation (3.1). Red dashed line shows the Fluctuation test's two sided critical value at 10%.

Figure 3.5 reports the results of the Fluctuation tests. Positive values of the (standardized) rMSFEs indicate that other core measures perform better than the standard exclusion index. Due to the shorter sample period, the timespan of the path of rMSFEs is the same as in the previous section: 1998:1 - 2011:6 for the two core CPI measures and 1991:12 - 2011:6 for trimmed PCE inflation.

Let me first focus on the results in Panels A and B on CPI inflation. The evidence suggests strong time-variation in the relative forecast performances with similar trimmed CPI and medium CPI performances. In the early 2000s, standard core measures had higher forecast accuracy. There was a reversal in relative forecast performance in the mid-2000s.

At this time, alternative measures performed better than exclusion core as indicated by the rMSFE exceeding the upper bound. Since the mid-2000s, however, all core CPI models have a similar forecast accuracy.

The first two observations are interesting insofar as all CPI core measures forecasted better than headline CPI from the beginning until the middle of the 2000s. Nevertheless, it appears that while CPIEXFE was able to pick up inflationary pressures around 2000 (housing prices), trimmed and median CPI inflation outperformed around 2005 (run-up of commodity prices). This suggests that there is no core measure that is strictly preferable to the other due to this time-variation. Instead, it might be useful to follow a set of core measures since it is impossible to predict which measure will do best in real-time.

When looking at PCE inflation (Panel C), the empirical patterns are different, but the conclusions appear to be in the same vein. In the first third of the sample and until the beginning of the 2000s, trimmed PCE inflation predicted significantly better than PCEExFE. In the middle of the sample, the difference in both models' forecast performance was not significant. With the onset of the Great Recession, standard exclusion PCE inflation forecasted better as indicated by the negative rMSFE. However, the difference in forecast performance is not statistically significant because the test statistic is within the bounds in recent years.

In recent years, the other core measures seem to perform worse than standard exclusion indices. This last observation is consistent with Ehrmann et al. (2018), who report results of Diebold and Mariano (1995) tests in subsamples from 2000-2018 for Europe. They find that the performance of other core measures worsened since 2007. They show that the reported RSME increased more than those of exclusion core for the same period.

To summarize, there is no clear answer to which core inflation measure is the best one. According to the analysis performed in this paper, the reason is that the relative forecast performances varied greatly over time. This explains why different analyses found different or ambiguous answers to the question. In reality, different measures perform better at different times. Since it is impossible to know which one is currently better in real-time, it is advisable to follow many indicators simultaneously. Another possibility could be to construct a composite index (see Cristadoro et al., 2005). However, in this section, I have shown that many core measures do not outperform headline inflation in recent years. Thus, combining them might not help to predict inflationary pressures. In the next section, I want to look at another alternative way of constructing a measure of core inflation.

3.4.2 *Average Inflation Indices*

> *If you want to predict inflation over the next three years, you really don't want to look just at inflation over the past 3 months or 6 months; you really want to look at inflation over the past three years. And if for whatever reason you want to use a shorter historical period [...] you should use core inflation, not headline.* (Krugman, 2011)

Another way to reduce volatility in the hope of decreasing noise, and, hence, another measure of underlying inflation is to increase sample selection size. This means looking at average inflation over the past 36 months instead of month-to-month changes. The underlying idea is that both transitory and persistent components drive the headline inflation. By taking averages of past inflation, one applies a one-sided moving average filter that filters out higher frequency fluctuations and should leave us with the low-frequency trend related to underlying inflation.[19]

[19] Alternative approaches are smoothed versions of headline inflation such as exponential smoothing (Cogley, 2002) or unobserved components-stochastic volatility (UC-SV) models (Stock and Watson, 2007). Another alternative is component-smoothed inflation measures (Gillitzer et al., 2006) such as the supercore measure of

By construction, this moving average places less weight on the current observation, which might result in a lagging indicator. This section concerns whether taking longer samples is indeed a good alternative to other core inflation measures.

A number of studies have explored the performance of averages of inflation measures (Blinder and Reis, 2005, Detmeister, 2011, or Bryan and Meyer, 2011). Among those, the consensus is that inflation rates for almost all measures predict future inflation better when averaged over a considerable number of months. To reconcile these findings in a framework accounting for instabilities, I perform Fluctuation tests on different sizes of sample intervals. In contrast to previous studies, I find that average inflation is not always a good measure of medium term inflation. On the one hand, averaging does reduce not only volatility but also eliminates signals. On the other hand, the relative forecasting performance of different sample lengths highly depends on the forecasting model considered. Therefore, the evidence suggests caveats in the use of averaging as a measure of core inflation.

To analyze different sample lengths of headline inflation, I compare year-on-year PCE headline inflation as a benchmark to both more low frequency and higher frequency PCE inflation rates. As forecasting model I will consider a modified version of equation (3.1) which follows Blinder and Reis (2005) and Crone et al. (2013).[20] Specifically, I allow for different sampling intervals in

$$\pi_{t,t+h} = \alpha + \theta x_{t-b,t} + \beta y_t + \varepsilon_t, \tag{3.7}$$

where $\pi_{t,t+h}$ is a measure of medium term inflation, $x_{t-b,t}$ is a predictive variable over the previous b months and α is a constant. Specifically,

underlying inflation of the ECB. It filters out the transitory component of components using econometric techniques. In particular, it selects those items that are estimated to co-move more with the business cycle.

[20] Model (3.1) is no longer adequate because considering polynomials of lagged average inflation would create filtered inflation itself.

h is the forecasting horizon, and b is the length of trend inflation. As predictors, I compare year-over-year PCE headline inflation to average PCE inflation for different sampling intervals, b. The variable y_t is a potential control variable. I will consider either (i) no control variable or (ii) the unemployment gap to consider a more structural Phillips curve model.

Accordingly, the test statistic of the Fluctuation test is modified in a similar way, comparing year-on-year headline ($b = 12$) to different average inflation rates

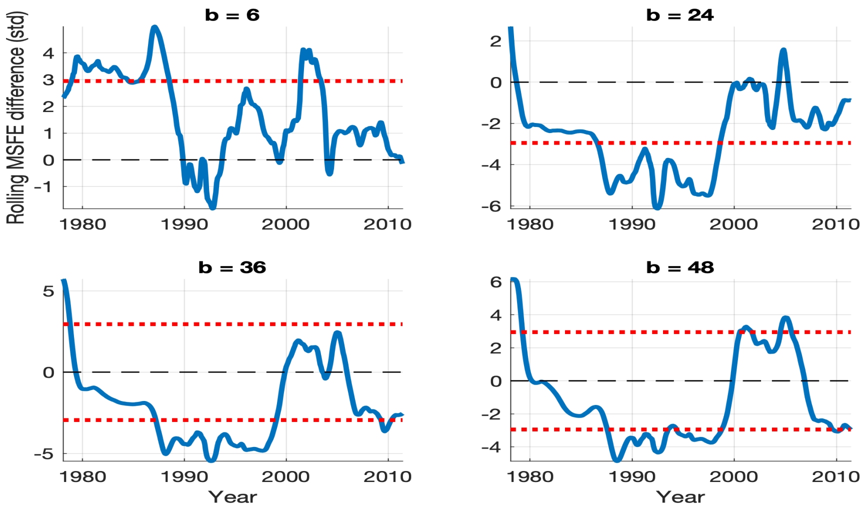
$$F_{t,m}^{OOS} = \widehat{\sigma}^{-1} m^{-1/2} \left(\sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{average})^2 - \sum_{j=t-m/2}^{j=t+m/2} (\widehat{\varepsilon}_{t+h}^{headline12})^2 \right), \tag{3.8}$$

where $\widehat{\varepsilon}_{t+h}^{average}$ and $\widehat{\varepsilon}_{t+h}^{headline12}$ are the forecast errors using either average inflation over different horizons, b, or year-on-year inflation as a predictor. Positive values of the test statistic represent a higher forecast accuracy of average inflation compared to the alternative measure.

A Longer Sample Interval Reduces Volatility But Also Signals. The idea of averaging is to reduce the influence of transitory fluctuations. Comparing the standard deviations of year-on-year headline (2.73), year-on-year core (2.47), and average headline inflation over the past 36 months (2.42) shows that the latter series successfully reduces volatility even more than core inflation.

Looking at the results from the Fluctuation tests, the outlook for improving predictive accuracy by averaging is rather pessimistic. Figure 3.6 displays the evidence for models with a different size of the coefficient b and without additional control variables. A positive rMSFE implies that the respective average inflation rate performs better than year-on-year inflation. In contrast, if the test statistic breaks through the lower bound, year-on-year headline inflation predicts better. Panels B

Figure 3.6: Comparison of Average Inflation Rates



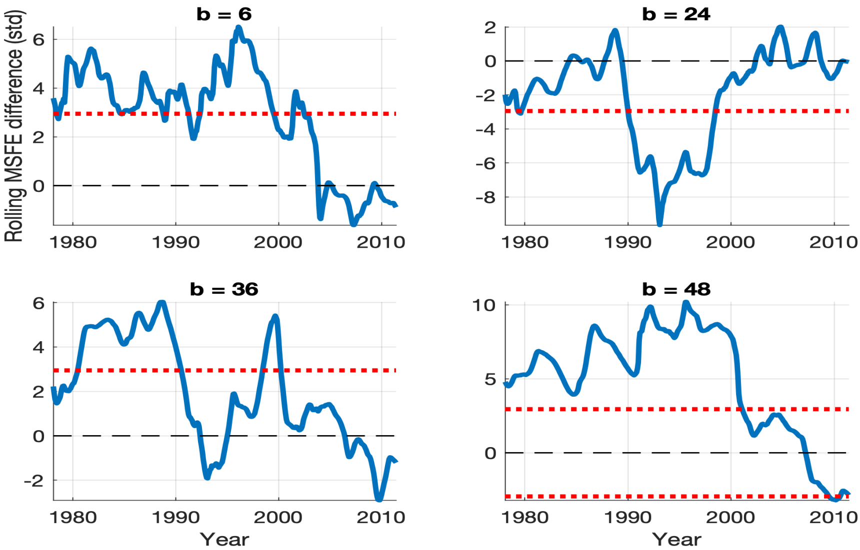
Note: Fluctuation Test comparing different average PCE inflation rates to year-on-year PCE inflation. Fluctuation test statistic, calculated as (standardized) difference between MSFE of the headline inflation and MSFE of the core model calculated over rolling windows ($m = 120$), across different specifications of equation (3.7). Red dashed line shows the Fluctuation test's two sided critical value at 10%.

to D show that averaging inflation improves predictive accuracy only at the beginning of the sample and in the early 2000s. The improvement is larger and statistically significant for larger averages, especially for average inflation over the past four years. In the same vein, inflation over the past six months ($b = 6$) appears to be performing better than year-on-year inflation for most of the time.

The Performance Depends on the Model. To study if this result is model-specific, I investigate a more structural model. To do so, I add as an additional control variable, y_t, the unemployment gap to equation (3.7). In this case, the forecast model resembles a Phillips curve relationship with the unemployment gap as a measure of economic activity.

Figure 3.7 illustrates the relative performance of different average

FIGURE 3.7: Comparison of Average Inflation Rates in PC Framework



Note: Fluctuation Test comparing different average PCE inflation rates to year-on-year PCE inflation. Fluctuation test statistic, calculated as (standardized) difference between MSFE of the headline inflation and MSFE of the core model calculated over rolling windows ($m = 120$), across different specifications of equation (3.7). Red dashed line shows the Fluctuation test's two sided critical value at 10%.

PCE inflation rates in this case. There is strong evidence of a non-linear relationship between the selected sample length and forecast performance. While six month and four-year average PCE inflation show better relative forecast performance than year-on-year inflation, the evidence for the other sample selection lengths is less strong. Irrespective of the horizon, the relationship breaks down at the beginning of the 2000s. Similar to all previous results, it appears that different measures of core inflation perform similarly in recent years.

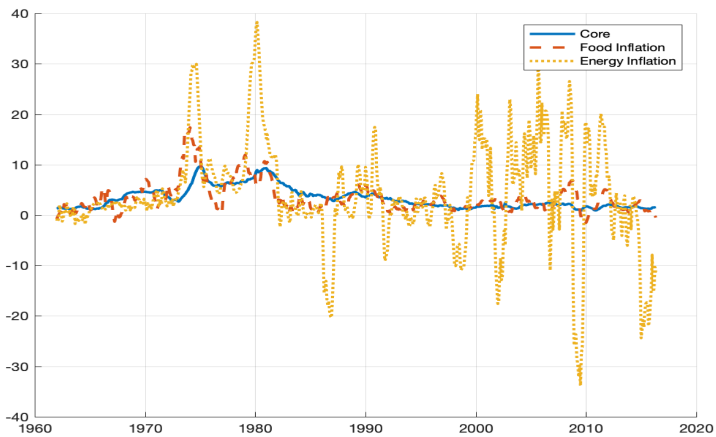
To summarize, the evidence of this exercise does not unambiguously confirm the hypothesis that longer sample periods provide better forecasts. Nevertheless, there are indeed periods and models when it might improve accuracy to take averages over longer periods. Notwithstanding, averaging comes with significant drawbacks. Those measures put less weight on the most recent monthly or quarterly data, assuming that it is so noisy that it has nothing useful to contribute to measuring underlying inflation. The results suggest that there are vital signals about future inflationary trends in recent data missed by averaging. The differences in results as compared to other studies originate from the consideration of instabilities and out-of-sample analysis.

3.5 Do non-core inflation components add information?

In the previous sections, I have studied if and when alternative core inflation measures have a forecasting edge towards headline inflation. The evidence suggests that this is indeed the case but only in the middle of the sample. This raises the question if the excluded items may contain important information about future inflationary pressures. In this case, one might have reduced volatility, but at the expense of eliminating signals. To test this possibility, this section revisits the empirical evidence on the informational content of Food and Energy prices.

Why should non-core components add not only noise but also information? First, the idea of core inflation is to exclude the most volatile items from headline inflation. Food and energy inflation indeed tend to be more volatile than core inflation, as illustrated by Figure 3.8. However, as outlined by Dolmas and Koenig (2019), not all items among the food and energy prices are the most volatile items. Their study reports that about 14 percent of food and energy items (by expenditure share) are less volatile than other core items. This means that we might involuntarily remove signals in the construction of core inflation and likewise allow some excess volatility to remain in the index. This concern becomes even more alarming when we contemplate that (i) the expenditure shares and (ii) volatilities of those items are time-varying.

Figure 3.8: Inflation Comparison: Non-Core Components of Inflation



Note: This figure compares the core and non-core components of PCE inflation. The figure considers year-on-year changes in core PCE inflation (solid blue line), food PCE inflation (red dashed) line and energy inflation (yellow dashed line).

Second, the purpose of measures of core inflation is to exclude transitory fluctuations (”blips”) in inflation. However, the implicit presumption is that food and energy inflation do not have a lasting impact on inflation. However, we have seen historical episodes where it can be argued that non-core inflation affected inflationary pressures/medium term inflation, e.g., during the oil prices shocks in the 1970s or during the run-up in global commodity prices in the mid-2000s. Both of these time periods can be seen in Figure 3.8. Another way in which non-core inflation might have a lasting effect on inflation is via indirect effects. Higher food and energy prices can influence other items in the consumer basket by affecting their production costs (production networks) or by changing inflation expectations (Coibion and Gorodnichenko, 2015). Accordingly, excluding non-core items might reduce the timeliness and reduce the reliability of signals of core inflation.

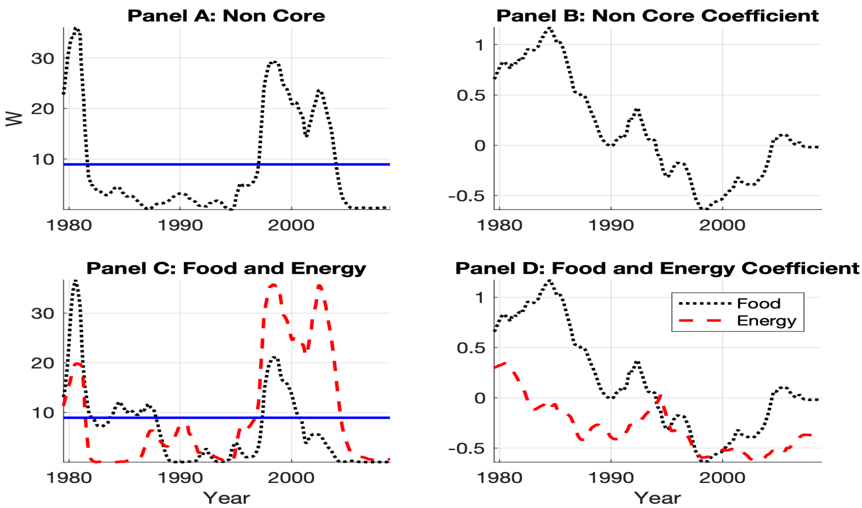
To assess whether non-core components add information to core forecasts, I consider the following regression:

$$\pi_{t,t+h} - \pi^{core}_{t+h|t} = \delta + \beta_i \pi^i_{t+h|t} + \beta_C x^{core}_{t+h|t} + \eta_{t+h}, \qquad (3.9)$$

where $\pi_{t,t+h}$ is actual medium term inflation and $\pi^i_{t+h|t}$ are forecasts generated from the model in equation (3.1) using either core, non-core, food or energy inflation as predictor, $x_{t,t-12}$. Forecasts from non-core inflation are useful beyond forecasts from core inflation if β_i is different from zero. One can also think about this regression as testing whether non-core forecasts add any marginal value to those forecasts derived from core inflation. This type of regression is usually used when testing the rationality of a forecast model (West and McCracken, 1998, or Mincer and Zarnowitz, 1969). Alterations of this framework are otherwise employed in studies that discuss the informational advantage of central banks with respect to private sector forecasts (Romer and Romer, 2000 or Hoesch et al., 2020).

In order to evaluate whether forecasts from non-core inflation and its components, $\pi^i_{t+h|t}$, where $i = \{noncore, food, energy\}$, provide additional information to core inflation's forecasts, I need to test if $\beta_i \neq 0$. Following the same reasoning as in Section 3.2, I want the test to be robust to the presence of possible time-variation in β_i. Therefore, I use the Fluctuation Rationality test proposed by Rossi and Sekhposyan (2016). The procedure is the following: A series of forecasts of medium term inflation using core and non-core components,$\pi^i_{t+h|t}$, and the forecast model (3.1) is generated. Then, equation (3.9) is estimated in rolling windows of size 180 months ($m = 180$) with estimates, $\widehat{\beta}_i$. For each rolling window estimate, I construct Wald-test statistics (W_i) of the null hypothesis of no information benefit, i.e., $\beta_i = 0$. The parameters are estimated by OLS, and HAC-robust standard errors (Newey and West, 1987) with a bandwidth equal to $P^{1/4}$ are constructed. The Fluctuation rationality test is a supremum test. Hence, the test statistic to test the information benefit, t_i , is the largest absolute value of the test statistic across all rolling windows. I use the tabulated critical values for model-free forecasts reported in Rossi and Sekhposyan (2016).[21] The time path of test statistics covers the time period from 1979:6-2008:6.

FIGURE 3.9: Information Advantage Test: PCE



Note: Fluctuation Rationality Test of equation (3.9). The figure reports the test statistic W (Wald test) for the null hypothesis $\beta_i \neq 0$ based on $m = 180$ at the 10% significance level.

Figure 3.9 shows the result of the information-advantage test for PCE inflation. In detail, it plots the Fluctuation-type test statistic (W_i) together with 5% critical values. The date on the horizontal axis provides information about the timing of the breakdowns by reflecting the center point of each rolling window.

I will focus first on the results of non-core inflation depicted in the two top panels. The evidence on the information advantage of non-core inflation complements the observations about forecast breakdowns of core inflation. Panel A of Figure 3.9 shows that non-core inflation added

[21]As discussed in Rossi and Sekhposyan (2016), using rolling window estimation with fixed window size guarantees that this assumption is satisfied, and the critical values in Table II can be used.

information to forecasts from core inflation in two time periods: (i) the beginning of the 1980s and (ii) around the 2000s. Both of these episodes are consistent with reversals in the forecast performance of core inflation. While the first can be linked to the period of high inflation in the 1970s, the second is related to the run-up of commodity prices in the mid-2000s. During that period, the increase was primarily driven by rapid economic growth in Asia. Rising global demand for commodities caused their prices to rise and put intense upward pressures on headline inflation. Interestingly, there appears to be no important information from non-core inflation in recent years.

In Panels C and D, I decompose the signaling channel of non-core inflation into its two components – (i) food (black dashed line) and (ii) energy inflation (red dashed line). The evidence points to the same two time periods where we can reject the hypothesis of no information benefit from food and energy inflation. Food and energy inflation provide no additional information in times when core inflation forecast performed well. The information Fluctuation test statistic is the largest absolute value of the test statistics across the rolling windows. For food inflation, this occurred in the 1980s.[22] In contrast, the signaling effect was strongest in the mid-2000s. Next, I turn to the coefficients. Both decrease over time and even turn negative. Negative coefficients likely reflect that core underpredicted headline inflation during the mid-2000s when increasing oil prices increased the aggregate inflation rate.

In summary, non-core components contain important information about future medium term inflation that can complement forecast from core inflation. This is particularly true in times when the relative forecast accuracy of core inflation becomes smaller. The evidence points to two periods where this happened. In these times, commodity prices had a lasting impact on medium term inflation. Thus, it would have

[22] Baumeister and Kilian (2014) show that there appears to be no evidence that oil price shocks can be associated with non-negligible increases in U.S. retail food prices in recent years.

yielded forecast benefits to include them in forecasting medium term inflation. This implies that there can be signals in non-core components, and removing them from core inflation might be detrimental. In recent years, however, the information in food and energy appears not to be the reason for the relative forecast performance of core inflation.

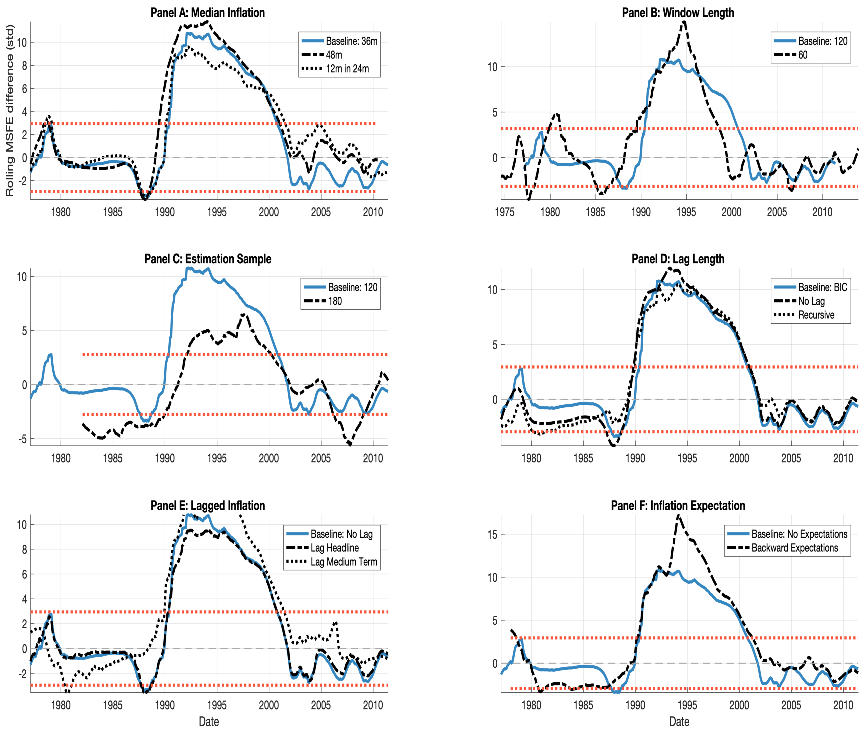
3.6 ROBUSTNESS ANALYSIS

In this Section, I test the robustness of the results from the previous Sections concerning different dimensions. In detail, I will look at the measure of medium term inflation, the choice of different parameters such as window length or estimation sample, or alternative forecasting models. The analysis of this section will focus on the results for PCE inflation, which are depicted in Panel B of Figure 3.3.

Other Measures of Medium Term Inflation. To address concerns about the correct measurement of medium term inflation, I consider alternative series for underlying inflation. Panel A of Figure 3.10 illustrates results for two alternatives: (i) inflation over the next 48 months, $\pi_{t,t+48}$, (solid blue line) and (ii) year-on-next-year inflation in 24 months, $\pi_{t+24,t+36}$, (red dashed line). From the evidence in Figure 3.10 it is clear that our main conclusions do not depend on the specific measure of medium term inflation.

Parameter Selection. I examine the robustness of the results when changing the size of the estimation sample, R, and the window size for forecast evaluation, m. In choosing these two, we face a trade-off between obtaining good estimates of the loss differences (larger m and R) and receiving a longer sample of relative rMSFEs to observe the evolution of relative forecast performance (smaller m and R). Panels B and C of Figure 3.10 compare our main results to different choices of m and R respectively. From Panel B, it is apparent that the main

Figure 3.10: Fluctuation Test: Robustness



Note: This Figure shows various robustness checks to the Fluctuation Test results of Section 3.3. In each Panel, the solid blue line shows the Fluctuation test statistic of Panel B in Figure 3.3. Whenever critical values vary, I plot those of the alternative model.

qualitative results remain unchanged when the window for the local losses is changed to 50, 100, or 150. The same is true when I choose different estimation sample sizes, e.g., 60 or 180.

Lag Length. The empirical results in Section 3.3 were based on a forecast model where the lag length was selected via a BIC criterion over the full sample period. The lag length was kept constant to minimize the impact of lag selection on forecasting performance for the comparison. I deal with this concern in two ways. First, I allow for a recursive lag-length selection, i.e., the lag length to be chosen each time the model is re-estimated. Second, I will abstract from lags of the explanatory variables in the forecasting model (3.1) and, instead, follow the forecast models in Blinder and Reis (2005) and Crone et al. (2013)

$$\pi_{t,t+h} = \alpha + \theta x_{t-12,t} + \varepsilon_t, \tag{3.10}$$

where $\pi_{t,t+h}$ is the measure of medium term inflation and $x_{t-12,t}$ is year-on-year change in the explanatory variable. Figure 3.10 shows that the results remain basically the same when the lag length is re-optimized each time the model is re-estimated.

Adding Lagged Headline Inflation. It is crucial to consider the robustness of the results to including autoregressive headline inflation terms. Stock and Watson (2007) reports that inflation in the 1970s was well described by an AR(1) process and Faust and Wright (2013) document that only surveys of inflation expectations can beat autoregressive models in forecasting headline inflation. Thus, I consider an alternative forecasting regression that includes either lagged headline inflation or lagged medium term inflation in

$$\pi_{t,t+h} = \alpha + \theta(L)x_{t-12,t} + \beta\pi_{t-l,t} + \varepsilon_t, \ t = 1, 2, ...T, \tag{3.11}$$

where $\pi_{t-l,t}$ is a measure of lagged inflation. I consider (i) lagged headline inflation, $l = 1$, or (ii) lagged medium term inflation, $l = 36$.

It is evident from Panel E of Figure 3.10 that our results are robust to adding autoregressive terms either in headline inflation (black dashed line) or medium term headline inflation (black dotted line).

Inflation Expectations. The most widely used model to predict inflation is based on the New-Keynesian Phillips curve. In this form, inflation is not only a function of economic activity but also of future inflation. To circumvent problems with the availability – most measures such as Michigan survey of households expectations or survey of professional forecaster are quarterly and start in 1984– we follow Blanchard (2016) or Ball and Mazumder (2011) and consider backward-looking inflation expectations. In detail, I augment the forecast model in (3.1) adding inflation expectations, π_t^e,

$$\pi_{t,t+h} = \alpha + \theta x_{t-12,t} + \beta \pi_t^e + \varepsilon_t, \tag{3.12}$$

where $\pi_t^e = 0.25(\pi_{t-1} + \pi_{t-2} + \pi_{t-3} + \pi_{t-4})$. This has two consequences. First, the model will better track the declining trend inflation rate in the 1980s and 1990s. Second, the model can flexibly account for increased inflation anchoring via changes in the window estimate $\widehat{\beta}$. Panel F of Figure 3.10 shows the results of the relative forecast performance of NK Phillips curves for PCE inflation. The evidence from this exercise does not qualitatively differ from our main results.

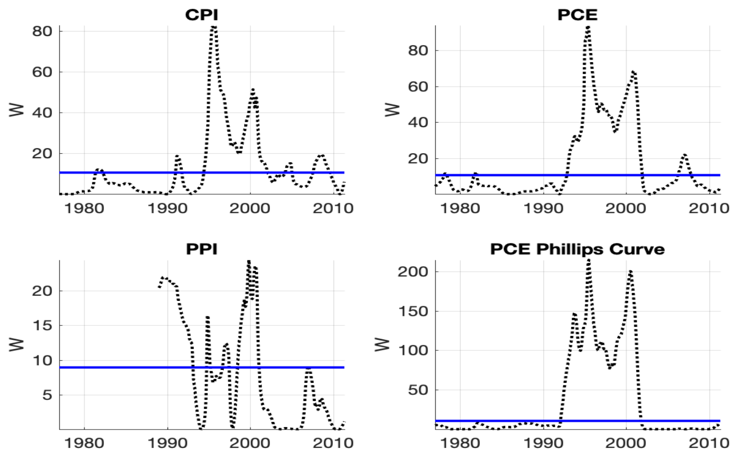
Marginal Forecast Value of Core Inflation. In his critique in the use of core inflation, Bullard (2011) argues that univariate models are the wrong metric to assess the usefulness of core inflation. Instead, the author argues that one should look at the marginal predictive value of core inflation in forecasting medium term inflation. For this purpose, one should consider a sophisticated model of inflation and then add core inflation. If the marginal value of adding core inflation is positive, then core inflation has some special information about future underlying inflation.

To address this concern, I apply a forecast rationality test similar to the Fluctuation-type test we performed in Section 3.5. In detail, I test if forecasts from core inflation can provide additional information to a forecasting model based on headline inflation. To assess whether core inflation adds marginal value to the forecasts based on headline inflation, consider the following regression:

$$\pi_{t+h} - \pi^H_{t+h|t} = \delta + \beta_c \pi^C_{t+h|t} + \beta_h \pi^H_{t+h|t} + \eta_{t+h}, \tag{3.13}$$

where $\pi^H_{t+h|t}$ and $\pi^C_{t+h|t}$ are the forecasts from model (3.1)with either headline or core inflation as predictive variable. Core forecasts have a marginal predictive value if $\beta_c \neq 0$. To test this type of questions in the presence of instabilities, I use the Fluctuation Rationality test proposed by Rossi and Sekhposyan (2016). The procedure and parameter choice is the same as in Section 3.5.

Figure 3.11: Test for Marginal Predictive Value of Core Inflation



Note: Fluctuation Rationality test of equation (3.13). This Figure reports the test statistic W for the null hypothesis $\beta_c \neq 0$ based on $m = 120$ at the 10% significance level.

Figure 3.11 sheds light on the marginal predictive value of core inflation. The Figure plots the Fluctuation-type test statistics over time, together with the critical values at the 10% level. It shows that there was indeed an information advantage of core inflation in the 1990s which, subsequently, deteriorated in the mid-2000s. However, the results of this exercise are not surprising because, under specific conditions, this test collapses to the Fluctuation test that compares the loss differences between each model separately.

3.7 Conclusion

This paper investigates the relative forecasting performance of different measures of core inflation over time. There are two central conclusions. First, I documented crucial instabilities in the forecast performance of core inflation measures. Second, the evidence in this study shows that no single measure of core inflation is the best predictor of underlying inflation in the U.S. over the whole history. Instead, I document instabilities in forecasting medium term inflation. I show that different core inflation indices perform better than autoregressive headline inflation models from the mid-1980s until the beginning of the 2000s. However, all considered models perform equally well in recent years. I leave for future work to investigate why different measures have no forecast advantage towards headline inflation in recent years.

When comparing the performance of different core inflation measures, I do not find that one measure of core dominates the others. Instead, I document that the relative performance is highly time-varying. In fact, they might offer different perspectives and insights that jointly help to understand developments in underlying inflation. As individually different core measures may not consistently give very precise or reliable signals, this calls for monitoring a wide range of measures of underlying inflation or constructing a composite index (Cristadoro et al., 2005, or Granziera and Sekhposyan, 2019).

BIBLIOGRAPHY

Acemoglu, D., Akcigit, U., and Kerr, W. (2016). Networks and the macroeconomy: An empirical exploration. *NBER Macroeconomics Annual*, 30:273–335.

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80:1977–2016.

Anand, R., Prasad, E. S., and Zhang, B. (2015). What measure of inflation should a developing country central bank target? *Journal of Monetary Economics*, 74(C):102–116.

Andrés, J., Ortega, E., and Vallés, J. (2008). Competition and inflation differentials in emu. *Journal of Economic Dynamics and Control*, 32(3):848–874.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61:821.

Aoki, K. (2001). Optimal monetary policy responses to relative-price changes. *Journal of Monetary Economics*, 48:55–80.

Ascari, G. and Sbordone, A. M. (2014). The macroeconomics of trend inflation. *Journal of Economic Literature*, 52:679–739.

Atalay, E. (2017). How important are sectoral shocks? *American Economic Journal: Macroeconomics*, 9:254–280.

Ball, L. and Mazumder, S. (2011). Inflation dynamics and the great recession. *Brookings Papers on Economic Activity*, pages 337–381.

Ball, L. and Romer, D. (1990). Real rigidities and the non-neutrality of money. *Review of Economic Studies*, 57:183–203.

Barnichon, R. and Mesters, G. (2020a). Identifying modern macro equations with old shocks. *The Quarterly Journal of Economics*, 135:2255–2298.

Barnichon, R. and Mesters, G. (2020b). The phillips multiplier. *Journal of Monetary Economics*, Forthcoming.

Barro, R. J. (1972). A theory of monopolistic price adjustment. *The Review of Economic Studies*, 39(1):17–26.

Basu, S. (1995). Intermediate goods and business cycles: Implications for productivity and welfare. *The American Economic Review*, 85:512–531.

Basu, S. (2005). Comment on: "implications of state-dependent pricing for dynamic macroeconomic modeling". *Journal of Monetary Economics*, 52:243–247.

Basu, S. (2019). Are price-cost markups rising in the united states? a discussion of the evidence. *Journal of Economic Perspectives*, 33:3–22.

Basu, S. and Leo, P. D. (2016). Should Central Banks Target Investment Prices? Boston College Working Papers in Economics 910, Boston College Department of Economics.

Baumeister, C. and Kilian, L. (2014). Do oil price increases cause higher food prices? *Economic Policy*, 29:691–747.

Baumeister, C. and Peersman, G. (2013). Time-varying effects of oil supply shocks on the us economy. *American Economic Journal: Macroeconomics*, 5:1–28.

Benigno, P. (2004). Optimal Monetary Policy in a Currency Area. *Journal of International Economics*, 63(2):293–320.

Bergin, P. R. and Corsetti, G. (2008). The extensive margin and monetary policy. *Journal of Monetary Economics*, 55(7):1222–1237.

Bernanke, B. S. (2007). Inflation expectations and inflation forecasting. *Monetary Economics Workshop of the NBER Summer Institute*.

Bernanke, B. S. (2010). Opening remarks : The economic outlook and monetary policy. *Economic Policy Symposium in Jackson Hole*.

Bigio, S. and La'O, J. (2020). Distortions in production networks. *The Quarterly Journal of Economics*, 135:2187–2253.

Bilbiie, F. O., Fujiwara, I., and Ghironi, F. (2014). Optimal monetary policy with endogenous entry and product variety. *Journal of Monetary Economics*, 64:1–20.

Bilbiie, F. O., Ghironi, F., and Melitz, M. J. (2008). Monetary Policy and Business Cycles with Endogenous Entry and Product Variety. In *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, pages 299–353. National Bureau of Economic Research, Inc.

Billi, R. M. and Gali, J. (2020). Gains from wage flexibility and the zero lower bound. *Oxford Bulletin of Economics and Statistics*, 82:1239–1261.

Bils, M. and Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112:947–985.

Bils, M., Klenow, P. J., and Malin, B. A. (2018). Resurrecting the role of the product market wedge in recessions. *American Economic Review*, 108:1118–1146.

Blanchard, O. (2016). The phillips curve: Back to the '60s? *American Economic Review*, 106:31–34.

Blanchard, O., Cerutti, E., and Summers, L. (2015). Inflation and activity - two explorations and their monetary policy implications. *IMF Working Papers*, 15.

Blanchard, O. and Simon, J. (2001). The long and large decline in u.s. output volatility. *Brookings Papers on Economic Activity*, pages 135–174.

Blanchard, O. J. and Gali, J. (2007). The macroeconomic effects of oil shocks: Why are the 2000s so different from the 1970s? *National Bureau of Economic Research no. 15467*.

Blanchard, O. J. and Galí, J. (2009). The macroeconomic effects of oil price shocks: Why are the 2000s so different from the 1970s? *International Dimensions of Monetary Policy*, pages 373–428.

Blinder, A. S. (1997). Commentary on " measuring short-run inflation for central bankers". *Review, Federal Reserve Bank of St. Louis*, 79.

Blinder, A. S. and Reis, R. (2005). Understanding the greenspan standard. *The Greenspan Era: Lessons for the Future*, 1196.

Bond, S., Hashemi, A., Kaplan, G., and Zoch, P. (2020). Some unpleasant markup arithmetic: Production function elasticities and their estimation from production data. *NBER Working Paper*, No. 27002.

Bouakez, H., Rachedi, O., and Santoro, E. (2018). Sectoral heterogeneity, production networks, and the effects of government spending. *HEC Working Papers*.

Bryan, M. F. and Cecchetti, S. G. (1994). Measuring core inflation. *Monetary policy*, pages 195–219.

Bryan, M. F. and Meyer, B. (2011). Should we even read the monthly inflation report? maybe not. then again.... *Atlanta Fed Macroblog*.

Bryan, M. F. and Pike, C. (1991). Median price changes: An alternative approach to measuring current monetary inflation. *Economic Commentary (Federal Reserve Bank of Cleveland)*, pages 1–4.

Bullard, J. (2011). Measuring inflation: The core is rotten. *Federal Reserve Bank of St. Louis Review*, 93:223–233.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12:383–398.

Campbell, J. R., Evans, C. L., Fisher, J. D., and Justiniano, A. (2012). Macroeconomic effects of federal reserve forward guidance. *Brookings Papers on Economic Activity*, pages 1–54.

Campbell, J. R., Fisher, J. D., Justiniano, A., and Melosi, L. (2016). Forward guidance and macroeconomic outcomes since the financial crisis. *NBER Macroeconomics Annual*, 31:283–357.

Carlstrom, C., Fuerst, T., Ghironi, F., and Hernandez, K. (2006). Relative price dynamics and the aggregate economy. *Manuscript, Boston*.

Carvalho, C. (2006). Heterogeneity in price stickiness and the real effects of monetary shocks. *Frontiers of Macroeconomics*, 2:1–58.

Carvalho, C. and Lee, J. W. (2011). Sectoral price facts in a sticky-price model. *Federal Reserve Bank of New York Staff Reports*.

Carvalho, C. and Schwartzman, F. (2015). Selection and monetary non-neutrality in time-dependent pricing models. *Journal of Monetary Economics*, 76:141–156.

Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, 28:23–48.

Carvalho, V. M. and Tahbaz-Salehi, A. (2019). Production networks: A primer. *Annual Review of Economics*, 11:635–663.

Christopoulou, R. and Vermeulen, P. (2012). Markups in the Euro area and the US over the period 1981–2004: a comparison of 50 sectors. *Empirical Economics*, 42(1):53–77.

Cienfuegos, N. C. (2019). The importance of production networks and sectoral heterogeneity for monetary policy. *SSRN Electronic Journal*.

Clark, T. E. (2001). Comparing measures of core inflation. *Economic Review-Federal Reserve Bank of Kansas City*, 86:5–32.

Clark, T. E. and McCracken, M. W. (2005). The power of tests of predictive ability in the presence of structural breaks. *Journal of Econometrics*, 124:1–31.

Cogley, T. (2002). A simple adaptive measure of core inflation. *Journal of Money, Credit, and Banking*, 34:94–113.

Coibion, O. and Gorodnichenko, Y. (2015). Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics*, 7:197–232.

Cooke, D. (2016). Optimal monetary policy with endogenous export participation. *Review of Economic Dynamics*, 21:72–88.

Cooper, R. and John, A. (1988). Coordinating coordination failures in keynesian models. *Quarterly Journal of Economics*, 103:50–55.

Cristadoro, R. R., Forni, M., Reichlin, L., and Veronese, G. (2005). A core inflation indicator for the euro area. *Journal of Money, Credit, and Banking*, 37.

Crone, T. M., Khettry, N. N. K., Mester, L. J., and Novak, J. A. (2013). Core measures of inflation as predictors of total inflation. *Journal of Money, Credit and Banking*, 45:539–560.

DeLoecker, J., Eeckhout, J., and Unger, G. (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics*, 135:561–644.

DeLoecker, J. and Warzynski, F. (2012). Markups and firm-level export status. *American Economic Review*, 102:2437–2471.

Detmeister, A. K. (2011). The usefulness of core pce inflation measures. *SSRN Electronic Journal*.

Detmeister, A. K. (2012). What should core inflation exclude? *SSRN Electronic Journal*.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13:253–263.

Dolmas, J. (2005). Trimmed mean pce inflation. *FRB Dallas working paper*.

Dolmas, J. and Koenig, E. F. (2019). Two measures of coreinflation: A comparison. *Federal Reserve of Dallas Working Papers Series*, 1903.

Edelstein, P. and Kilian, L. (2009). How sensitive are consumer expenditures to retail energy prices? *Journal of Monetary Economics*, 56:766–779.

Ehrmann, M., Ferrucci, G., Lenza, M., and O'brien, D. (2018). Measures of underlying inflation for the euro area. *ECB Economic Bulletin*, 1:94–115.

Erceg, C. J., Henderson, D. W., and Levin, A. T. (2000). Optimal monetary policy with staggered wage and price contracts. *Journal of Monetary Economics*, 46(2):281–313.

Etro, F. and Rossi, L. (2015). New-keynesian phillips curve with bertrand competition and endogenous entry. *Journal of Economic Dynamics and Control*, 51:318–340.

Eusepi, S., Hobijn, B., and Tambalotti, A. (2011). Condi: A cost-of-nominal-distortions index. *American Economic Journal: Macroeconomics*, 3:53–91.

Faia, E. (2012). Oligopolistic competition and optimal monetary policy. *Journal of Economic Dynamics and Control*, 36(11):1760–1774.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. *Handbook of Economic Forecasting*, 2:2–56.

Fischer, S. (1977). Long-term contracts, rational expectations, and the optimal money supply rule. *Journal of Political Economy*, 85:191–205.

Galesi, A. and Rachedi, O. (2019). Services deepening and the transmission of monetary policy. *Journal of the European Economic Association*, 17:1261–1293.

Gali, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*, volume 2.

Gali, J. (2020). The effects of a money-financed fiscal stimulus. *Journal of Monetary Economics*, 115:1–19.

Gali, J. and Gambetti, L. (2019). Has the u.s. wage phillips curve flattened? a semi-structural exploration. *CEPR Discussion Papers.*

Giacomini, R. and Rossi, B. (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics*, 25:595–620.

Giacomini, R. and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, 74:1545–1578.

Giannone, D., Lenza, M., Momferatou, D., and Onorante, L. (2014). Short-term inflation projections: A bayesian vector autoregressive approach. *International Journal of Forecasting*, 30:635–644.

Gillitzer, C., Simon, J., et al. (2006). Component-smoothed inflation: Estimating the persistent component of inflation in real time. *Reserve Bank of Australia, Working Paper*.

Golosov, M. and Lucas, R. E. (2007). Menu costs and phillips curves. *Journal of Political Economy*, 115(2):171–199.

Gordon, R. J. (1975). Alternative responses of policy to external supply shocks. *Brookings Papers on Economic Activity*, 1.

Gorodnichenko, Y. and Weber, M. (2016a). Are sticky prices costly?: Evidence from the stock market. *American Economic Review*, 106:165–199.

Gorodnichenko, Y. and Weber, M. (2016b). Are sticky prices costly? evidence from the stock market. *American Economic Review*, 106(1):165–99.

Granziera, E. and Sekhposyan, T. (2019). Predicting relative forecasting performance: An empirical investigation. *International Journal of Forecasting*, 35:1636–1657.

Hall, R. E. (1986). Market structure and macroeconomic fluctuations. *Brookings Papers on Economic Activity*, 1986:285–322.

Hall, R. E. (1988). The relation between price and marginal cost in u.s. industry. *Journal of Political Economy*, 96:921–947.

Hazell, J., Herreño, J., Nakamura, E., Steinsson, J., et al. (2020). The slope of the phillips curve: evidence from us states. *Unpublished manuscript*.

Herrera, A. M. and Pesavento, E. (2009). Oil price shocks, systematic monetary policy, and the "great moderation". *Macroeconomic Dynamics*, 13:107–137.

Hoesch, L., Rossi, B., and Sekhposyan, T. (2020). Has the information channel of monetary policy disappeared? revisiting the empirical evidence. *Federal Reserve Bank of San Francisco, Working Paper Series*, pages 1–43.

Hooper, P., Mishkin, F. S., and Sufi, A. (2020). Prospects for inflation in a high pressure economy: Is the phillips curve dead or is it just hibernating? *Research in Economics*, 74:26–62.

Horvath, M. (2000). Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45:69–106.

Huang, K. X. and Liu, Z. (2005). Inflation targeting: What inflation rate to target? *Journal of Monetary Economics*, 52(8):1435–1462.

Hulten, C. R. (1978). Growth accounting with intermediate inputs. *Review of Economic Studies*, 45:511–518.

Imbs, J., Jondeau, E., and Pelgrin, F. (2011). Sectoral phillips curves and the aggregate phillips curve. *Journal of Monetary Economics*, 58:328–344.

Jorda, O., Marti, C., Nechio, F., Tallman, E., et al. (2019). Why is inflation low globally? *FRBSF Economic Letter*, 19:15.

Katayama, M. and Kim, K. H. (2018). Intersectoral labor immobility, sectoral comovement, and news shocks. *Journal of Money, Credit and Banking*, 50:77–114.

Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18:39–43.

Kiley, M. T. (2015). Low inflation in the united states : A summary of recent research. *FEDS Notes, Washington: Board of Governors of the Federal Reserve System*, 2015.

Kimball, M. S. (1995). The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking*, 27:1241–1277.

Klenow, P. J. and Kryvtsov, O. (2008). State-dependent or time-dependent pricing: Does it matter for recent u.s. inflation? *Quarterly Journal of Economics*, 123:863–904.

Krugman, P. (2011). Core madness. *New York Times*.

Lee, D. and Wolpin, K. I. (2006). Intersectoral labor mobility and the growth of the service sector. *Econometrica*, 74:1–46.

Lenza, M. and Reichlin, L. (2011). Should the ecb use core inflation as a signal for medium-term inflation? *VoxEU.org*.

Lewis, V. (2013). Optimal monetary policy and firm entry. *Macroeconomic Dynamics*, 17(8):1687–1710.

Luciani, M. and Trezzi, R. (2019). Comparing two measures of core inflation: Pce excluding food and energy vs. the trimmed mean pce index. *FEDS Notes*, 2019.

Mankiw, N. G. and Reis, R. (2003). What Measure of Inflation Should a Central Bank Target? *Journal of the European Economic Association*, 1(5):1058–1086.

Mavroeidis, S., Plagborg-Møller, M., and Stock, J. H. (2014). Empirical evidence on inflation expectations in the new keynesian phillips curve. *Journal of Economic Literature*, 52:124–188.

McLeay, M. and Tenreyro, S. (2020). Optimal inflation and the identification of the phillips curve. *NBER Macroeconomics Annual*, 34:4–255.

Mincer, J. and Zarnowitz, V. (1969). The evaluation of economic forecasts. *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER New York.

Mishkin, F. S. (2007). Headline versus core inflation in the conduct of monetary policy. *Business Cycles, International Transmission and Macroeconomic Policies Conference, HEC Montreal, Montreal, Canada.*

Nakamura, E. and Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. *Quarterly Journal of Economics*, 123:1415–1464.

Nakamura, E. and Steinsson, J. (2010). Monetary non-neutrality in a multisector menu cost model. *Quarterly Journal of Economics*, 125:961–1013.

Nason, J. M. (2006). Instability in us inflation: 1967-2005. *Economic Review-Federal Reserve Bank of Atlanta*, 91:39.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708.

of Economic Advisers, C. (2016). Benefits of competition and indicators of market power. *Issue Brief.*

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64:1263–1297.

Pasten, E., Schoenle, R., and Weber, M. (2019). The propagation of monetary policy shocks in a heterogeneous production economy. *Journal of Monetary Economics.*

Pasten, E., Schoenle, R., and Weber, M. (2020). Price rigidities and the granular origins of aggregate fluctuations. *Working Papers Central Bank of Chile.*

Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861–19571. *Economica*, 25:283–299.

Powell, J. (2019). Congressional testimony on july 10 2019 (c-span).

Ramey, V. A. and Vine, D. J. (2010). Oil, automobiles, and the u.s. economy: How much have things really changed? *NBER Macroeconomics Annual*, 25:334–367.

Reinelt, T. and Meier, M. (2020). Monetary policy, markup dispersion, and aggregate tfp. *ECB Working Paper*.

Rich, R. W. and Steindel, C. (2005). A review of core inflation and an evaluation of its measures. *Federal Reserve Bank of New York Staff Reports*, 236.

Rich, R. W. and Steindel, C. (2007). A comparison of measures of core inflation. *FRBNY Economic Policy Review*, pages 19–38.

Romer, C. D. and Romer, D. H. (2000). Federal reserve information and the behavior of interest rates. *American Economic Review*, 90:429–457.

Rossi, B. (2013). Advances in forecasting under instability. *Handbook of Economic Forecasting*, 2:1203–1324.

Rossi, B. and Sekhposyan, T. (2010). Have economic models' forecasting performance for us output growth and inflation changed over time, and when? *International Journal of Forecasting*, 26:808–835.

Rossi, B. and Sekhposyan, T. (2016). Forecast rationality tests in the presence of instabilities, with applications to federal reserve and survey forecasts. *Journal of Applied Econometrics*, 31:507–532.

Rotemberg, J. J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER macroeconomics annual*, 12:297–346.

Rotemberg, J. J. and Woodford, M. (1999). Interest rate rules in an estimated sticky price model. In *Monetary policy rules*, pages 57–126. University of Chicago Press.

Rowe, N. (2011). No, atlanta and st. louis fed, you can't test whether core is useful that way. *Blog: A Worthwhile Canadian Initiative*.

Rubbo, E. (2020). Networks, phillips curves and monetary policy. Technical report, Working paper, Harvard University.

Samuelson, P. A. and Solow, R. M. (1960). Analytical aspects of anti-inflation policy. *American Economic Review*, 50:177–194.

Sheshinski, E. and Weiss, Y. (1977). Inflation and costs of price adjustment. *The Review of Economic Studies*, 44(2):287–303.

Smaghi, L. B. (2011). Ignoring the core can keep inflation at bay. *Financial Times*, 1.

Stock, J. and Watson, M. (2019). Slack and cyclically sensitive inflation. *National Bureau of Economic Research Working Paper*, 25987.

Stock, J. H. and Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44:293–335.

Stock, J. H. and Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41:788–829.

Stock, J. H. and Watson, M. W. (2007). Why has u.s. inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39:3–33.

Stock, J. H. and Watson, M. W. (2013). Phillips curve inflation forecasts. *Understanding Inflation and the Implications for Monetary Policy*, pages 101–186.

Taylor, J. B. (1980). Aggregate dynamics and staggered contracts. *Journal of Political Economy*, 88:1–23.

Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Confer. Series on Public Policy*, 39:195–214.

Thornton, D. L. (2011). Core versus headline inflation: An opportunity for greater transparency. *Economic Synopses*, 2011.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica*, 64:1067–1084.

West, K. D. and McCracken, M. W. (1998). Regression-based tests of predictive ability. *International Economic Review*, 39:817–840.

Woodford, M. (2002). Inflation stabilization and welfare. *Contributions in Macroeconomics*, 2(1).

Woodford, M. (2011). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Yellen, J. L. (2015). Inflation dynamics and monetary policy. *Philip Gamble Memorial Lecture, University of Massachusetts, Amherst, Massachusetts*.