Keynesian Fluctuations:
Nominal Rigidity or Frictional Coordination?*

George-Marios Angeletos
MIT and NBER

November 10, 2017

Abstract

I revisit the notion that recessions are driven by low aggregate demand. I first explain why the underlying partial-equilibrium logic is misleading and why this notion finds little place in the RBC framework. I next review the assumptions that accommodate it within the New Keynesian framework and some of the ensuing challenges. I finally discuss an alternative approach, which sums up my current research agenda. This approach helps understand recessions as the product of a novel type of coordination failure, one that does not require multiple equilibria. It anchors expectations of future outcomes to past outcomes, dulls the general-equilibrium effects of innovations in the underlying fundamentals, and makes room for forces akin to animal spirits. It can serve as a substitute for the dominant formalization of demand-driven fluctuations, shifting the focus from nominal rigidity to frictional coordination; but it can also enrich the New Keynesian framework, modifying its predictions in manners that seem both conceptually appealing and empirically relevant.

*This article is based on the Schumpeter Lecture that I gave at the 2016 Annual Meeting of the European Economic Association, in Geneva. I thank Dirk Krueger, Harris Dellas and Fabrice Collard for detailed feedback; Fabrizio Zilibotti for early encouragement; and Chen Lian and Karthik Sastry for assistance. I am also grateful to my co-authors in the line of research upon which a large part of this lecture/article is based. Email: angelet@mit.edu.
1 Introduction

Understanding the role of aggregate demand in business cycles is central to macroeconomics. For instance, a drop in aggregate demand triggered by the financial crisis and ensuing collapse in housing prices is considered to be one of the main, if not the main, force behind the Great Recession and the slow recovery from it (Mian, Rao, and Sufi, 2013; Mian and Sufi, 2014; Hall, 2011; Eggertsson and Krugman, 2012).

This is in line with the Keynesian narrative, namely, that recessions are caused by low, or deficient, aggregate demand. In this article, I start by explaining why this narrative is subtle, why it finds little place in the neoclassical/RBC framework, and how exactly it is accommodated in the New Keynesian framework. I next review how my own research has sought to shed new light on the theoretical underpinnings of this narrative, on its empirical content, and on its policy implications.

But let me first clarify the philosophy behind this article. As Robert E. Lucas, Jr. once said,¹

“Theoretical economists ... do not ask for words that ‘explain’ what equations mean. We ask for equations that explain what words mean.”

In this article, I take for granted the validity of the Keynesian narrative, but ask for the equations that explain the words.

In Subsection 2.1, I start by pointing out two elementary points, which are a recurring theme of this article. The first is that, while the aforementioned narrative may seem obvious from a partial-equilibrium (PE) perspective, it can be misleading when considering the general-equilibrium (GE) response to economy-wide shocks.² The second is that the notion of aggregate demand in modern macroeconomic theory is best understood as a particular type of relative demand in the Arrow-Debreu framework, namely, the demand for goods today relative to the demand for goods tomorrow.

These points help distinguish modern macroeconomics from the old IS-LM framework, which lacked micro-foundations, rest on fussy extrapolations from PE reasoning, made ad hoc assumptions directly about behavior (e.g., consumption functions) as opposed to constraints and incentives, and failed to provide a satisfactory understanding on what aggregate demand means and how it matters. They also explain the theoretical constructs that macroeconomists often use as proxies for drops in aggregate demand, such as negative shocks to the subjective discount factor of the consumers.

Such shocks should not be taken literally. They help capture within simple models the effects that more realistic shocks have in more elaborate models, such as the effects of a credit crunch on the consumers (Eggertsson and Krugman, 2012; Guerrieri and Lorenzoni, 2017), of higher precautionary savings (Challe et al., 2017), or of pessimistic beliefs about future income (Lorenzoni, 2009). Accord-

¹Nobel Lecture delivered at Trinity College in 2001.
²Henceforth, PE and GE as acronyms for, respectively, partial equilibrium and general equilibrium.
ing to the Keynesian narrative, any of these forces translates into a drop in aggregate demand and triggers a recession, that is, a joint reduction in employment, output, consumption and investment.

Building on these points, Subsection 2.2 reviews why this notion is at odds with the neoclassical, or Real Business Cycle (RBC), framework. By reducing the demand for goods today relative to goods tomorrow, a negative discount-rate shock triggers a drop in consumer spending. But it also depresses the price of goods today relative to goods tomorrow, which in turn encourages the firms to invest more. Furthermore, insofar as the households have an urge to consume less of all goods today (as when they face tighter credit constraints), they also have an incentive to work more. As a result, the drop in consumer spending is accompanied by a boom in employment, output, and investment.

Although it may be tempting to discard this prediction as empirically implausible, there is an important lesson to take home. The main limitation of the baseline RBC model is also its main strength: it assumes no departure from the Arrow-Debreu framework, no market failure, and seemingly nothing more than the elementary assumptions that underlie the familiar demand-and-supply curves in microeconomics.\textsuperscript{3} It follows that the inconsistency of the Keynesian narrative with the baseline RBC model is proof of how subtle, and potentially misleading, this narrative is. It may appear self-evident from a PE perspective, but it falls apart in the most basic GE context.

To salvage the Keynesian narrative, one must therefore move, not only beyond PE reasoning, but also beyond the most basic GE setting. One must figure out the hidden ingredients (i.e., the additional assumptions) that enable the narrative to work.\textsuperscript{4}

In the New Keynesian framework, the key ingredients are two. First, there is a constraint in how nominal prices (or nominal wages) adjust to aggregate shocks; in macro jargon, “prices are sticky” or there is “nominal rigidity”. Second, the policy maker is unable to render this constraint non-operative; in macro jargon, “monetary policy does not replicate flexible prices”. I fill in the equations that explain the words in Subsection 2.4.

The aforementioned assumptions may be realistic, but they are not part of the demand-and-supply picture that is familiar from microeconomics. They also lead to three empirical challenges. First, the model is able to generate realistic business cycles with demand shocks (that is, fluctuations with the

\textsuperscript{3}By saying “seemingly”, I anticipate the following point. The easily recognizable assumptions of the RBC framework, which are familiar from microeconomics, are price-taking behavior, individual optimization, rational expectations, and market clearing. A more delicate assumption, which my work relaxes, is that of common knowledge of the current state of the economy and of its future prospects.

\textsuperscript{4}Early attempts to reconcile the Keynesian narrative with micro-founded, GE settings included a literature on coordination failures and sunspot fluctuations, which I discuss in the sequel, and the works of Lucas (1972, 1973) and Barro (1976, 1978) on rational nominal confusion and unexpected monetary shocks. These works, which built on the Friedman’s notion of the “natural rate” (Friedman, 1968), sough to explain why an apparent trade off between unemployment and inflation—a Philips curve—could be present in the data and, yet, could not be exploited by monetary policy in a systematic way. By contrast, the New Keynesian framework, which eventually dominated over the alternatives and which I am concerned with in this article, assumes that there is an actual, and exploitable, structural relation between inflation and real economic activity, in the form of the New Keynesian Phillips curve. I discuss the central position that this relation plays in the theory and the discomfiting evidence about it in Subsection 2.5.
right co-movement among the real quantities) only by combining the aforementioned assumptions with more dubious “bells and whistles” (more on this later). Second, menu-cost models that explain the microeconomic data on prices suggest that the macro-level nominal rigidity may be quite smaller than what the New Keynesian framework requires in order to match the effects of identified monetary shocks and to produce plausible business cycles. Last but not least, the model predicts that the gap of the equilibrium outcomes from their flexible-price (or RBC) counterparts ought to be manifested in inflation, a prediction that is hard to reconcile with the data.\(^5\)

In addition to these empirical challenges, the New Keynesian framework suffers, at least in my view, from an inherent conceptual defect: it ties the notion of demand-driven fluctuations with those of nominal rigidity and monetary non-neutrality. In other words, the only way one is allowed to make sense of demand-driven fluctuations within that framework is by equating a drop in aggregate demand with, effectively, a monetary contraction.\(^6\)

This connection does not seem to be present in the data. It is also at odds with a Keynesian tradition that emphasized the idea that low aggregate demand is the product of a coordination failure and that sought to capture this idea in models with multiple equilibria and sunspot fluctuations.\(^7\)

I therefore find it appealing, if not imperative, to try to relax the aforementioned restriction: while I recognize the potential empirical importance of nominal rigidities and monetary non-neutrality, I would like to be able to make sense of the idea that the Great Recession and the slow recovery from it could have been driven in large part by a “deficiency in aggregate demand” even when ignoring nominal rigidities and the zero lower bound on monetary policy.\(^8\)

Furthermore, the empirical problems of the New Keynesian framework are tightly connected to the aforementioned restriction. For instance, because that framework equates a deficiency in aggregate demand with, effectively, a monetary contraction, it also implies, counterfactually, that the Great Recession should have been the Great Deflation. Relaxing the aforementioned restriction is therefore, not only theoretically appealing, but also empirically promising.

In a series of papers, I have thus advocated a new formalization of demand-driven fluctuations, one that shifts the focus from nominal rigidities to imperfect coordination. By the latter, I do not mean a coordination failure in the sense of multiple equilibria. Instead, I mean the accommodation of the possibility that the agents may have an incomplete, and idiosyncratic, understanding of one another’s behavior, of the current state of the economy, and of its future prospects.

---

\(^5\)This relates to the empirical failings of the New Keynesian Philips Curve, which I briefly discuss in the end of Section 4.

\(^6\)By the latter I mean a contraction relative to the benchmark of replicating flexible-price allocations. This in turn could be so because monetary policy is suboptimal, because it cannot tell apart different shocks, or because of the zero lower bound on interest rates has become binding. See Subsection 2.4 for further details.

\(^7\)Diamond (1982); Cass and Shell (1983); Cooper and John (1988); Woodford (1991); Guesnerie and Woodford (1993); Benhabib and Farmer (1994, 1999).

\(^8\)As already mentioned, the older literature on coordination failures and sunspot fluctuations also tried rationalize demand-driven fluctuations without nominal rigidities. Recent works aimed at the same goal include Beaudry and Portier (2013) and Benhabib, Wang, and Wen (2015).
From a game-theoretic perspective, this friction amounts to dropping the conventional assumption that the underlying state of Nature is common knowledge and, thereby, accommodating higher-order uncertainty, that is, uncertainty about the beliefs and actions of others. Such uncertainty can be the product of either the segmentation of market interactions and the decentralization of the information, along the lines of Lucas (1972), or of rational inattention and costly contemplation, along the lines of Sims (2003, 2010) and Tirole (2015). I follow the first modeling approach in Angeletos and La’O (2013) and Angeletos and Lian (2017), the second in Angeletos and La’O (2017) and Angeletos and Sastry (2017); I consider both of them in Angeletos and Lian (2016a). One way or another, what is key is that the agents are unable to reach perfect consensus about the current state of the economy and its future prospects. This lack of consensus represents a coordination friction.

This friction has the flavor of a coordination failure. But whereas the traditional formalization of coordination failure operates across different equilibria, the one formalized in my work operates along any given equilibrium. It can therefore be embedded in models featuring a unique equilibrium, including the most basic versions of the RBC and New Keynesian frameworks.

In Section 3, I review how this friction helps make sense of the Keynesian narrative within the RBC framework. This is accomplished in two distinct but complementary ways. In the one (Angeletos and La’O, 2013; Angeletos, Collard, and Dellas, 2015), I introduce a new kind of shocks, which cause exogenous shifts in the higher-order beliefs of the underlying fundamentals and which ultimately rationalize waves of pessimism and optimism about the short-term economic outlook. In the other (Angeletos and Lian, 2016c), I refrain from the introduction of such shocks and, instead, show how the considered friction modifies the propagation of familiar forms of demand shocks.

Either way, the Keynesian narrative finds a new place in the theory, one that does not rest on the presence of significant nominal rigidities and on the failure, or unwillingness, of monetary policy to replicate flexible prices. By the same token, the focus is shifted from constraints on monetary policy, such as the zero lower bound, to the difficulty agents face in digesting what's going on in the economy and in coordinating their beliefs and actions. A new perspective therefore emerges about the Great Recession. Perhaps the kind of demand shocks documented in Mian and Sufi (2014) contributed to a recession for different reasons than those presumed in the standard model (namely, sticky prices and the zero lower bound on monetary policy). And perhaps this also explains why the Great Recession was not the Great Deflation.\footnote{Coibion and Gorodnichenko (2015b) seek to reconcile the facts with the theory by documenting that expectations of inflation, as measured in surveys, did not turn negative during the Great Recession. But the theory predicts that expectations of inflation should have turned negative, so this evidence does not really solve the puzzle; in my view, it only transfers it from one dimension to another. For a complete explanation within the New Keynesian framework, which attributes the absence of deflation to offsetting cost-push shocks, see Christiano, Eichenbaum, and Trabandt (2015).}

In these regards, my approach can be viewed as a substitute for the New Keynesian framework. But it can also serve as a powerful complement to it. This is true, not only in terms of offering a

\footnote{Coibion and Gorodnichenko (2015b) seek to reconcile the facts with the theory by documenting that expectations of inflation, as measured in surveys, did not turn negative during the Great Recession. But the theory predicts that expectations of inflation should have turned negative, so this evidence does not really solve the puzzle; in my view, it only transfers it from one dimension to another. For a complete explanation within the New Keynesian framework, which attributes the absence of deflation to offsetting cost-push shocks, see Christiano, Eichenbaum, and Trabandt (2015).}
complementary formalization of demand-driven fluctuations (along the lines described above), but also in terms of modifying its policy predictions in empirically appealing ways and offering a microfoundations for some of its more dubious bells and whistles. This is the theme of Section 4.

To be concrete, consider the response of the economy to news about future monetary policy (“forward guidance”) during a liquidity trap. In the New Keynesian framework, such news have large effects on current outcomes because they trigger large shifts in the expectations that consumers and firms form about the behavior of others and, thereby, about aggregate spending and inflation.

These predictions are at odds with the available empirical evidence, a challenge known in the literature as the “forward guidance puzzle” (Del Negro, Giannoni, and Patterson, 2012; McKay, Nakamura, and Steinsson, 2016). But a variant of this puzzle appears to apply more generally: the macroeconomic times series indicate small and sluggish responses of a variety of variables to a variety of identified shocks.\(^{10}\) What is more, the sluggishness in the response of the actual outcomes appears to be accompanied by an even more pronounced sluggishness in the response of expectations (Coibion and Gorodnichenko, 2015a).

My approach offers a parsimonious explanation to these salient features of the data. The prevailing paradigm assumes that the underlying shocks, or news, and their likely effects on economic outcomes such as inflation and income are common knowledge. Once this assumption is relaxed, expectations of future outcomes become anchored to past outcomes. This is because the agents lack confidence that the other agents will adjust their beliefs and behavior in response to the available news. As a result, the economy behaves as if the agents were myopic in the sense of discounting the influence of the underlying shocks on future economic outcomes (Angeletos and Lian, 2016a). This lessens the forward guidance puzzle, offers a rationale for the front loading of fiscal stimuli, and slows down the recovery of the economy from a recession once “good news” start arriving.

In addition, the considered friction causes the economy to behave as if the agents were backward-looking and were pegging their current choices on past outcomes. This property is the manifestation of the gradual adjustment in higher-order beliefs over time. As a result, my approach provides a microfoundation of some of the more dubious bells and whistles that the New Keynesian framework has relied on in order to match the macroeconomic time series, such as adjustment costs in investment and the hybrid version of the New Keynesian Philips curve (Angeletos and Huo, 2017).

Last but not least, because expectations of future outcomes work through GE mechanisms, anchoring the former is akin to dulling the latter (Angeletos and Lian, 2016a, 2017). This helps reduce the distance between the predictions of fully-fledged macroeconomic models and the underlying PE intuitions—perhaps the simplistic, PE logic about demand and supply is not as misleading after all.

\(^{10}\)For instance, Gali (1999) documents a sluggish response of real quantities to identified technology shocks, while Christiano, Eichenbaum, and Evans (2005) document a sluggish response of both the real quantities and inflation to identified monetary shocks. See also the complementary discussion in Sims (2003).
Remark 1. Informational frictions can have important effects even in a single-agent decision context. For instance, rational inattention can dull the response of individual choices to idiosyncratic shocks. This article, however, is focused on the additional, and distinct, effects that arise in multiple-agent contexts from the interaction of information frictions with GE mechanisms, or with strategic complementarity. This interaction is at the core of the literatures on global games and beauty contests. For a synthesis of these literatures, for extensive references, and for a more thorough exposition of the modeling role of higher-order uncertainty in macroeconomics and finance, I refer the interested reader to my chapter in the Handbook of Macroeconomics (Angeletos and Lian, 2016b).

Remark 2. Informational frictions can rationalize monetary non-neutrality even in the absence of sticky prices and menu costs. This is an important point, which goes back to Lucas (1972, 1973) and Barro (1976, 1978) and has been revisited more recently by Mankiw and Reis (2002), Woodford (2003a), and Mackowiak and Wiederholt (2009). This, however, is not the theme of this article.  

2 Aggregate Demand and the Business Cycle: It’s Complicated

In this section, I review a few elementary lessons from modern (post-IS-LM) macroeconomics that are relevant for my purposes. I first explain the difficulty in extrapolating from familiar, PE intuitions about demand and supply to their GE counterparts. I next show that the Keynesian narrative finds no place within an elementary, two-period, GE model, which is essentially a stripped down version of the RBC framework. I finally review how this notion is formalized within the New Keynesian framework and discuss some of the challenges with this formalization.

2.1 Back to the Basics

The most familiar figure in economics is probably the one that illustrates the demand for, and the supply of, an arbitrary good. Letting \( q \) denote the quantity of that good and \( p \) its price, we can express its demand and its supply as, respectively,

\[
q = D(p, X_d) \quad \text{and} \quad q = S(p, X_s),
\]

where \( X = (X_d, X_s) \) are all the other factors that affect demand and supply, such as the prices of all other goods, the consumers’ tastes and income, and the firms’ technologies. Partial equilibrium is defined by equating demand and supply, holding \( X \) constant. This is represented by the intersection of the two solid lines in Figure 1.

In this PE context, a negative demand shock means a leftward shift in the demand curve, \( D \), induced by a change in some of the factors in \( X \), holding the supply curve, \( S \), constant. Assuming

\[11\] Another important issue that I abstract from in this article is the role of financial frictions.
that the former is downward sloping and the latter is upward sloping, both the equilibrium quantity and the equilibrium price fall.

This is, of course, trivial. But how can one extrapolate from this picture to understand what macroeconomists mean when they say that business cycles are driven by aggregate demand shocks? It might be tempting to do this by changing notation from lower-case variables, $q$ and $p$; to upper-case variables, $Q$ and $P$. But one has to be careful what the latter mean.

When we talk about demand and supply at the micro level, there are always some other goods and some other markets in the background: the little $p$ in Figure 1 is the price of one good relative to that of all other goods in the economy. If the big $Q$ were an index of all the goods in a Arrow-Debreu economy, then the big $P$ would have to be the price of that index relative to itself, which is tautologically one. To have a meaningful extrapolation from Figure 1 to a macroeconomic context, it therefore has to be that $Q$ captures only some of the goods in an Arrow-Debreu economy and that $P$ is their price relative to the rest of the goods.

In the light of this elementary point, the convention in modern macroeconomics is to consider multi-period settings, to emphasize forward-looking behavior, and to interpret $Q$ as an index of the goods produced and consumed today and $P$ as their price relative to the goods produced and consumed tomorrow. By the same token, a shock to aggregate demand means a shock in the demand for goods today relative to the demand for goods. Note that this differs from the old IS-LM framework, which abstracted from forward-looking behavior and interpreted $P$ as the nominal price level.\footnote{The New Keynesian framework has nevertheless reduced the distance between the micro-founded perspective described above and the IS-LM framework by connecting the real interest rate to the inflation rate. Also, Lucas (1972, 1973) and Barro (1976, 1978) sought to make sense of a demand-and-supply picture at the aggregate level by interpreting the variable in the vertical axis as the gap between the actual and the perceived, or expected, nominal price level.}

How should we think about a “drop in aggregate demand” from this modern, micro-founded perspective? To be concrete, consider the drop in consumer spending during the Great Recession. The evidence in Mian, Rao, and Sufi (2013) and Mian and Sufi (2014) suggests that this drop was

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Demand Shocks in Partial Equilibrium}
\end{figure}
triggered by a sharp tightening in consumer credit and by deleveraging. For the present purposes, however, it suffices to abstract from the precise trigger of a drop in consumer spending and, instead, proxy it with an exogenous discount-rate shock, that is, a shock to inter-temporal preferences. By causing a sudden “urge” to consume less today relative to tomorrow, a negative discount-rate shock can mimic, at least qualitatively, the effect of tighter consumer credit on consumer spending.\footnote{The use of discount-rate shocks as proxies for shocks to consumer credit is quite common in the macroeconomics literature. See, for example, Eggertsson and Woodford (2003), Werning (2012) and Christiano, Eichenbaum, and Trabandt (2015). For more realistic treatments, see Eggertsson and Krugman (2012), Guerrieri and Lorenzoni (2017) and\footnote{A minimalistic GE model of this kind can also be found in Barro (1997).}}

Let us take such a shock for granted and let us ask how this propagates in the economy, how it affects all the macroeconomic quantities of interest. To answer this question, it is useful to consider the minimal GE model for the job. This model has a representative household, which means that I abstract from heterogeneity. It has two periods, “today” and “tomorrow,” and three goods, leisure today, consumption today, and consumption tomorrow.\footnote{A minimalistic GE model of this kind can also be found in Barro (1997).}

Clearly, these modeling choices are only a very rough approximation of the real world. They are designed to isolate what is of essence for the narrative under consideration from a number of bells-and-whistles often contained in richer macroeconomic models. Having two periods is of essence so that we can talk meaningful about a shock to “aggregate demand,” that is, a preference shock shifting the demand for goods today relative to tomorrow. It is also necessary for accommodating investment. Finally, having a consumption-leisure choice in the first period allows us to endogenize labor supply and to study the effects of the shock on employment, while abstracting from such a choice in the second choice is only for simplicity.

I next study two versions of this simple model. In the first, nominal prices are “flexible” and monetary policy is neutral; this can be thought of as a stripped-down version of the RBC framework. In the second, nominal prices are “sticky” and monetary policy is non-neutral; this represents a stripped-down version of the New Keynesian framework.

### 2.2 Demand Shocks in the RBC Framework

I now set up the neoclassical, or RBC, version of my minimalist model. There is a representative household, living two periods, \( t \in \{1, 2\} \). Her preferences are given by

\[
\theta U(c_1, \ell_1) + U(c_2, \ell_2),
\]

where \( U \) and \( u \) are strictly increasing and strictly concave, \( c_t \) and \( \ell_t \) denote, respectively, consumption and leisure in period \( t \), for \( t \in \{1, 2\} \), and \( \theta \) is an exogenous preference parameter that determines how much the household values goods today vs goods tomorrow. Without serious loss of generality, I assume that preferences are separable between consumption and leisure, so that \( U(c, \ell) = u(c) + v(\ell) \),
where $u$ and $v$ strictly increasing and strictly concave. I also impose $\ell_2 = \bar{\ell}_2$, for some exogenous $\bar{\ell}_2$, so that can I concentrate on labor supply and employment fluctuations in the first period.

Consider next the production side of the economy. In each period, a representative firm uses capital and labor to produce the final good according to a Cobb-Douglas technology. Output in periods 1 and 2 is therefore given by, respectively,

$$y_1 = AF (k_1, n_1) \quad \text{and} \quad y_2 = F (k_2, n_2), \quad (2)$$

where $k_t$ and $n_t$ are the amounts of capital and labor in period $t$, $F(k, n) = k^{\alpha}n^{1-\alpha}$, $\alpha \in (0, 1)$, and $A$ is an exogenous parameter that measures the total factor productivity (TFP) in the first period.

To close the model, note that, in the first period, the household can either consume her income or save it into capital to be used in the second period. But since the second period is the last period of the household’s life, the household will consume all its income and all accumulated capital in that period. Assuming that depreciation is zero, we can thus write the resource constraint of the economy in, respectively, periods 1 and 2 as follows:

$$c_1 + k_2 = y_1 + k_1 \quad \text{and} \quad c_2 = y_2 + k_2, \quad (3)$$

These conditions represent also market clearing in the product markets. The labor markets, on the other hand, clear if and only if

$$n_1 = 1 - \ell_1 \quad \text{and} \quad n_2 = 1 - \bar{\ell}_2, \quad (4)$$

where the total endowment of time has been normalized to 1. Finally, let $i_1 \equiv k_2 - k_1$ denote the investment in the first period.

The model described above allows one to generate variation in the equilibrium outcomes by introducing variation in $\theta$ and $A$. If $\theta$ falls, the representative household wants to consume less today relative to tomorrow. In this sense, $\theta$ represents an “aggregate demand shock.” By contrast, $A$ represents a “aggregate supply shock.” The question addressed in the rest of this section is how the first-period equilibrium outcomes respond to each of these shocks.

As long as the First Welfare Theorem applies (which it does under standard assumptions), we can understand the competitive equilibrium and address the aforementioned question by solving a simple planning problem: that of maximizing (1) subject to (2)-(4). The optimality conditions of this problem,

---

15 As standard in macro, I assume household owns the capital and rents it to the firm; with complete markets, it does not matter whether the capital is owned by the household or the firm.

16 I am abstracting from news and noise shocks, that is, shocks to beliefs of tomorrow’s productivity (Beaudry and Portier, 2006). Such shocks, too, are unable to generate realistic business cycles in either the fully-fledged RBC model or the two-period variant studied here. This explains why the literature has combined such shocks with exotic features in preferences (Jaimovich and Rebelo, 2009) or a departure from flexible prices (Lorenzoni, 2009; Barsky and Sims, 2011).
and the prices that support the planner’s solution as part of a competitive equilibrium, are given by
the following:

\[ \frac{v'(\ell_1)}{u'(c_1)} = w = F_n(k_1, n_1) \tag{5} \]

\[ \frac{\theta u'(c_1)}{u'(c_2)} = 1 + r = 1 + F_k(k_2, n_2) \tag{6} \]

where \( w \) denotes the real wage in the first period and \( r \) denotes the real interest rate between the two
periods (this is the big \( P \) in our earlier discussion).\(^{17}\)

These conditions have a familiar interpretation from microeconomics. The planner equates the
marginal rates of substitution (MRS) of any pair of commodities, be it the consumption-leisure bundle
in the first period or the consumption levels in the two periods, with the corresponding marginal rates
of transformation (MRT). The competitive equilibrium replicates the solution to the planning problem
by having the household equate the MRS of any two commodities with their relative price and the
firms equate the latter with the corresponding MRTs.

The solution to the system of equations (3)-(6) pins down the planner’s optimum or, equivalently,
the equilibrium. By investigating how this solution varies as we vary \( \theta \) and \( A \), we can then shed light
on the macroeconomic effects of, respectively, demand and supply shocks.\(^{18}\)

Consider first a negative supply shock, that is, a drop in \( A \). It is easy to show that this triggers
a drop in employment, output, and investment at \( t = 1 \). The intuition is simple. Because today’s
productivity decreases, the substitution effect between work and leisure drives down employment
and output, which in turn justifies a reduction in both consumption and investment. This reviews the
RBC model, in which negative technology shocks cause recessions.

What about a negative “demand shock,” that is, a drop in \( \theta \)? When \( \theta \) falls, consumers have “an
urge to save,” so consumption today goes down. But employment, output, and investment go up!

Why is this happening? From the perspective of the planner, the drop in the desire to consume
today frees up resources for investment, even if we hold employment and output fixed. Moreover, the
reduction in \( \theta \) means a drop in the relative demand for all the goods consumed today, including leisure.
This stimulates employment and output. All in all, a drop in consumer spending is accompanied by
a boom in employment, output, and investment, not by a recession.\(^{19}\)

\(^{17}\)The second condition does not contain an expectation operator because I have abstracted from uncertainty.

\(^{18}\)Strictly speaking, I am only conducting comparative statics with respect to the parameters \( \theta \) and \( A \). But the insights
developed here directly extend to the impulse responses of the fully-fledged RBC model with respect to productivity and
discount-factor shocks.

\(^{19}\)Note that the argument rests on modeling the demand shock as a shift in inter-temporal preferences, holding constant the
intra-temporal preferences between consumption and leisure. A shock to the latter kind of preferences represents, instead, a
shock to the supply of labor. Such a shock can generate positive co-movement between employment, output, consumption
and investment, but it does not offer a compelling explanation of the observed business cycles. (The Great Recession was
not the Great Vacation.) Note, however, that such a shock is formally equivalent to a shock in the labor wedge and that this
wedge appears to be highly cyclical in the data. This underscores why any theory that aspires to improve upon the RBC
model must ultimately explain the observed cyclicality of the labor wedge (Chari, Kehoe, and McGrattan, 2007).
The “magic” that translates this property from the planner’s solution to the competitive equilibrium lies in the adjustment of two relative prices, the real interest rate and the real wage. As consumers try to spend less on goods today, the real interest rate—which is the relative price of these goods—falls. This stimulates the demand for investment. The shock therefore causes a drop in one component of the demand for goods today and an increase in another. Finally, as the household tries to consume less leisure today, the real wage—which is the relative price of leisure—falls. This encourages the firms to employ more workers and produce more goods today, which means that the supply of goods actually increases.

This illustrates why the Keynesian narrative finds no place in the baseline RBC model: within that model, business cycles have to be attributed to supply shocks.\textsuperscript{20} To make sense of the narrative, one can not just build on the most basic principles of microeconomics. Instead, one must also allow for something additional, something more subtle.

In the New Keynesian framework, which I review next, this extra ingredient is the combination of nominal rigidity with certain “mistakes,” or constraints, in the conduct of monetary policy. In the alternative that I favor, it is a certain imperfection in how well the agents understand what’s going on in the economy, modeled as lack of common knowledge and higher-order uncertainty.

2.3 Parenthesis: Demand Shocks and the Labor Wedge

Before filling in the details, let me clarify the following point. It is possible to generate, within the RBC model, positive co-movement out of demand shocks if one transforms the latter to movements in either TFP or the labor wedge.

To understand what we mean by the first scenario, consider the model described above and suppose that \( A \) happens to be an increasing function of \( \theta \). Then, provided that a drop in \( \theta \) comes together with a sufficiently large drop in \( A \), it is clearly possible that a drop in \( \theta \) triggers a recession. This is essentially route taken by Bai, Rios-Rull, and Storesletten (2017): that paper develops an extension of the RBC model that adds search frictions in commodity markets and that lets preference shocks generate endogenous movements in measured TFP.

To understand what we mean by the second scenario, note first the labor wedge is defined as the gap between the marginal product of labor and the MRS between consumption and leisure. In the model described above, this gap is zero. To accommodate a non-zero labor wedge, we must thus replace the planner’s intra-temporal condition with the following variant:

\[
\frac{v'(l_1)}{w'(c_1)} = (1 - \tau_1) AF_n(k_1, n_1),
\]

\textsuperscript{20}Although I have focused on consumer-specific demand shocks, essentially the same insights apply to investment-specific demand shocks: shocks that shift the expectations of future productivity and the return to current investment fail to generate realistic business cycles within the baseline RBC model.
where \( \tau_1 \neq 0 \). Taken literally, \( \tau_1 \) can be interpreted as a tax on labor demand or labor supply. More broadly, it captures any distortion that drives the relevant MRS and MRT apart. Note then that, holding both \( \theta \) and \( A \) constant, an increase in \( \tau_1 \) generates a joint drop in employment, output, consumption and investment.\(^{21}\) It follows that any modification of the considered model that lets a demand shock to act, in effect, as an increase in \( \tau_1 \) will do the job.

Both the New Keynesian framework and my proposed alternative can be understood under these lenses: they let a drop in \( \theta \), or a drop in “consumer confidence” (properly defined), trigger an increase in the realized labor wedge.\(^{22}\) But is there a good reason to prefer this kind of approach to the one that lets the demand shock become a technology shock? Yes. The bulk of the employment fluctuations in the data is orthogonal to the fluctuations in TFP and labor productivity. It follows that an empirically successful theory of demand-driven fluctuations is, not one that transforms the demand shock to a technology shock, but rather one that transforms the demand shock to a labor-wedge shock.\(^{23}\)

### 2.4 Demand Shocks in the New Keynesian Framework

The New Keynesian framework departs from the baseline RBC framework by adding monopoly power and, most importantly, sticky prices: firms are price-setters, rather than price-takers, and adjust their nominal prices only periodically.\(^{24}\) This makes monetary policy non-neutral. As a result, we can no longer answer the question of how the economy responds to an aggregate demand shock without specifying how monetary policy itself responds to that shock.

There is, however, an important policy benchmark that sheds light on how the model works more generally. This corresponds to a monetary policy that “replicates flexible prices,” that is, one that implements the same real outcomes as those that would have obtained in the absence of nominal rigidity. Insofar as monetary policy does not have to substitute for missing tax instruments, such a policy is actually optimal (Correia, Nicolini, and Teles, 2008). But even when such a policy is suboptimal, or just infeasible, the benchmark defined by it is instrumental for understanding whether and how the New Keynesian model can accommodate the desired narrative.

When monetary policy replicates flexible prices, the New Keynesian model reduces, in effect, to the RBC model. It follows that the New Keynesian framework is unable to make sense of the desired narrative unless monetary policy deviates from that benchmark.

---

\(^{21}\)In a PE context, a tax on labor can have offsetting substitution and income effects. In the GE context under consideration, however, \( \tau_1 \) has only a substitution effect: it drives a wedge between MRS and MRT without affecting the resource constraint of the economy. This explains why an increase in \( \tau_1 \) unambiguously reduces \( n_1, y_1, c_1 \) and \( i_1 (= k_2) \).

\(^{22}\)Note that the labor wedge, as defined above, combines the wedge between the real wage and the marginal product of labor (what is the markup in the New Keynesian framework) and the wedge between the real wage and the MRS.

\(^{23}\)This is also a key challenge of a recent literature that generates business cycles out of “uncertainty shocks” by letting such shocks trigger endogenous movements in aggregate TFP and labor productivity (Bloom, 2009; Bloom et al., 2012), as well as of an earlier literature that attempted to do the same with sunspots (Benhabib and Farmer, 1999).

\(^{24}\)The monopoly distortion is irrelevant for the mechanics of the model and can be negated by an appropriate subsidy. The key departure from RBC is the nominal rigidity.
How can such deviation help the model accommodate the desired narrative? By letting the exogenous drop in consumer spending transform, in effect, to a negative monetary shock. By this I mean that monetary policy has to contract relative to the aforementioned benchmark.

Let me elaborate. Because of the monopoly power and the nominal rigidity, the equilibrium of the economy no more coincides with the planner’s solution studied in Subsection 2.2. Accordingly, conditions (5) and (6) no more hold. Instead, the following variants hold:

\[
\frac{v' (c_1)}{u' (c_1)} = w = \frac{1}{\mu_1} AF_n(k_1, n_1) \tag{8}
\]

\[
\frac{\theta u' (c_1)}{u' (c_2)} = 1 + r = 1 + \frac{1}{\mu_2} F_k(k_2, n_2) \tag{9}
\]

where \(\mu_t\) denotes (one plus) the realized monopoly markup in period \(t\).²⁵

Comparing conditions (8) and (9) to conditions (5) and (6), we see that the only difference is the emergence of the markup \(\mu_t\) as a wedge between the relevant MRSs and MRTs. Note, in particular, that an increase in \(\mu_t\) is isomorphic to a joint tax on labor and capital. This explains why \(\mu_1\) plays exactly the same role as the labor wedge in Subsection 2.3, and why \(\mu_2\) shows up as an “Euler wedge” between the two periods.

The intuition is familiar from microeconomics: monopoly distortions are akin to tax distortions. What is important for our purposes, however, is that the New Keynesian framework lets the realized monopoly distortion, \(\mu_t\), and the associated labor and Euler wedges be under the control of monetary policy. This is where macroeconomics departs from microeconomics.

When prices are flexible, \(\mu_t\) can higher than one, reflecting the monopoly distortion, but is exogenous to monetary policy. Let \(\mu^*_t\) denote the equilibrium value of the markup that obtains under flexible prices (the “ideal” markup). A natural starting point is to assume that \(\mu^*_t\) is time- and state-invariant. I adopt this assumption here in order to simplify the exposition: \(\mu^*_1 = \mu^*_2 = \mu^*\), for some \(\mu^* > 1\) that is invariant to \(\theta\) and \(A\). It then follows that, although the equilibrium may feature a lower level of employment and output than the planner’s solution due to the monopoly distortion, it shares essentially the same comparative statics with respect to either \(A\) or \(\theta\).

When, instead, prices are sticky, the realized markup hinges on whether monetary policy adheres to or deviates from the aforementioned policy benchmark. If monetary policy replicates flexible prices, the firms produce their ideal amount of output, the realized marginal cost equals the realized marginal

²⁵The New Keynesian framework features two additional equilibrium conditions, which are not central for the arguments made in this subsection but are essential for understanding how monetary policy can control the equilibrium allocations and the associated markups by varying its policy instrument. The one is the Euler condition for the nominal bond or, equivalently, the Fischer equation: the real interest rate is equated to the nominal one net of the inflation rate. The other condition is a structural equation that related inflation to real economic activity, namely the New Keynesian Philips curve. By combining these two conditions with conditions (8) and (9), one then sees how monetary policy can control jointly the realized markups, the associated allocation, and the inflation rate by varying the nominal interest rate.
revenue, and the realized markup is just right (i.e., $\mu_t = \mu^*$). If, instead, monetary policy is expansionary relative to that benchmark, the firms end up producing too much, the realized marginal cost exceeds marginal revenue, and the realized markup is too low (i.e., $\mu_t < \mu^*$). And if monetary policy is contractionary relative to that benchmark, the firms end up producing too little and the realized markup is too high (i.e., $\mu_t > \mu^*$).

Of course, the monetary authority controls $\mu_t$ only indirectly: by varying the nominal interest rate.26 To illustrate, let me momentarily shut down investment, so that $c_t = y_t$ and condition (9) is dropped. Suppose further that production is linear in labor, so that $F(k, n) = n = 1 - \ell$. Suppose further that prices are completely rigid in the first period but flexible in the second. This means that the second-period allocation is invariant to monetary policy, that $\mu_2 = \mu^*$, and that the second-period levels of consumption and output are given by $c_2^* = y_2^* = 1 - \bar{\ell}_2$, where $\bar{\ell}_2$ is the exogenously specified amount of leisure. Then, conditions (8) and (9) can be written as follows:

$$\frac{v'(1 - \frac{1}{A}c_1)}{u'(c_1)} = \frac{1}{\mu_1} A \quad \text{and} \quad \frac{\theta u'(c_1)}{u'(c_2^*)} = 1 + r$$

where I used the facts that $c_1 = y_1 = An_1$ and $n_1 = 1 - \ell_1$ to replace $v'(\ell_1)$ in the first condition with $v'(1 - \frac{1}{A}c_1)$. Finally, suppose that the monetary authority guarantees that the inflation rate in the second period is zero, which means that $r$, the real interest rate between the two periods, moves one-to-one the nominal interest rate. It is then evident that, by varying the latter, the monetary authority can vary $r$ and, thereby, also vary $c_1$ and $\mu_1$.27,28

While this example is special, the logic applies more generally. Within the New Keynesian model, the nominal rigidity plays a dual role. On the one hand, it shapes how real allocations respond to shocks for a given monetary policy rule. On the other hand, it allows monetary policy to manipulate real allocations and to serve as a substitute for taxation or market regulation: it is as if the realized markup is a policy instrument under the control of the monetary authority.

---

26In an older literature, the policy instrument is money supply. This required the specification of a demand for money. Following Woodford (2003b), the modern practice focuses on the “cashless limit” in which monetary policy affects real allocations and welfare only by manipulating the nominal interest rate.

27I am abstracting here from the Zero Lower Bound on the nominal interest rate, which translates to a lower bound on the value of $\mu_1$. Also note that, for expositional reasons, I have swept under the carpet the delicate issue of how exactly the monetary authority guarantees that the inflation rate between the two periods is zero. In the infinite-horizon version of the New Keynesian model, it is easy to show that there exists an equilibrium in which the monetary authority induces zero inflation—and replicates the flexible-price outcomes—for all periods after the first one ($t \geq 2$) by following an appropriate interest-rate policy from $t = 2$ and on. It then follows that, along this equilibrium, the real interest rate between $t = 1$ and $t = 2$ moves one-to-one with the nominal interest rate, exactly as assumed here. There is a more delicate issue of whether this particular equilibrium is the unique one, but I will not deal with this issue here.

28Clearly, there is a specific value for $r$, denoted here by $r^*$, that induces $\mu_1$ to coincide with $\mu_1^*$. This is the so-called “natural rate of interest”: it is the one that replicates the underlying flexible-price allocation. If $r$ is lower than $r^*$, the induced $\mu_1$ is lower than $\mu^*$ and, equivalently, $c_1$ is higher than its flexible-price counterpart. The converse is true if $r$ is higher than $r^*$. A monetary policy that induces the real interest rate to be below (respectively, above) the “natural rate” therefore induces an expansion (respectively, a contraction) relative to the flexible-price benchmark. The magnitude of this expansion (respectively, contraction) is a monotone transformation of the gap between $\mu_1$ and $\mu^*$. 
Let me now explain how this feature of the New Keynesian model helps accommodate the Keynesian narrative. Switch on investment and consider, once again, a drop in $\theta$. As already noted, such a negative demand shock triggers a boom in employment ($n_1$), output ($y_1$), and investment ($i_1$) when the nominal rigidity is absent (as in the RBC setting of Subsection 2.2) or, equivalently, when monetary policy replicates flexible prices. Note that this boom is associated with a drop in $r^*$, the natural rate: as the demand for goods today falls, their relative price also falls.

Suppose, now, that monetary policy fails to replicate flexible prices and, instead, lets the current markup, $\mu_1$, increase as $\theta$ falls. This happens when the monetary authority does not allow the actual interest rate, $r$, to fall as much as the natural rate, $r^*$. By doing so, the monetary authority triggers a contraction relative to the underlying flexible-price outcomes. It is therefore as if the monopoly distortion has been intensified or a tax has been imposed on of firm sales. Other things equal, such a policy causes $n_1$ and $y_1$ to fall. It follows that the increase in $\mu_1$ has exactly the opposite effect on employment and output than the drop in $\theta$ has when prices are flexible. To the extent that the increase in $\mu_1$ is sufficiently large, the overall effect on employment, output, and even investment can be negative.$^{29}$

2.5 Some Challenges

Let me review the lessons learned so far. In the RBC model, which is arguably the most elementary GE model one can think of, a negative demand shock generates a boom rather than a recession. The New Keynesian model shares this “pathological” prediction if monetary policy replicates flexible-price allocations (which, under certain conditions, is actually the optimal thing to do). To accommodate the more plausible scenario that a negative demand shock triggers a recession, the New Keynesian model has to assume that monetary policy causes a contraction relative to the aforementioned benchmark. In this sense, the New Keynesian model generates the desirable prediction only by transforming the underlying demand shock into a contractionary monetary-policy shock.

Is this mechanism empirically plausible? The answer to this question depends on whether one focuses on the prediction that negative demand shocks trigger recessions or on the assumptions that lead to this prediction. Throughout this article, I take for granted that this prediction is the “right” one (i.e., consistent with the facts). What I want to quibble about is the assumptions that allow this prediction to obtain within the model and an additional prediction that follows from these assumptions and is actually hard to reconcile with the data.

Let me first focus on the assumptions. The most obvious one is that prices are sticky or, more broadly, that there is some kind of nominal rigidity. This assumption is certainly consistent with the

$^{29}$For investment to fall with the drop in $\theta$, it has to be that the associated increase in $\mu_1$ is even larger than the one required for employment and output to fall, or that it is accompanied by an increase in $\mu_2$. That is, either the current contraction in monetary policy has to be sufficiently severe, or the contraction has to be sufficiently persistent.
available micro-economic evidence (Klenow and Malin, 2010). Yet, even if there is significant nominal rigidity at the micro level, its bite at the macro level can be small because of the reason first highlighted in Caplin and Spulber (1987). When firms must pay a fixed cost (a “menu cost”) in order to adjust their prices, they will opt to adjust only infrequently; but they will also move their prices by a relatively large amount whenever they adjust. As a result, the average adjustment in the price level following an aggregate shock can be comparable to the one that would have obtained in the absence of nominal rigidity, even if most of the firms don’t adjust prices most of the time. In a nutshell, monetary policy could be neutral at the aggregate level even if there is significant rigidity at the micro level.

Although Caplin and Spulber’s result rested on strong assumptions, it highlighted how the presence of significant nominal rigidity at the micro level could be consistent with ample price flexibility at the macro level. See Caballero and Engel (1993a,b, 2007) for the robustness of this insight and for a thorough analysis of the aggregate implications of price rigidity. These early works also paved the way to a more recent literature, which address the relevant quantitative question: is the macro-level rigidity that the New Keynesian model requires in order to match the macroeconomic time series consistent with the one predicted by menu-cost models when the latter are calibrated to match key moments of the available micro data, such as the frequency and size of price adjustments at the firm level?

A first, and rather discomforting, answer to this question was provided in an influential paper by Golosov and Lucas (2007): the macro-level nominal rigidity that could be justified on the basis of the available microeconomic evidence coupled with a basic menu-cost model was shown to be significantly smaller than the one assumed in textbook calibrations of the New Keynesian model and even smaller than the one assumed in estimated DSGE models such as Christiano, Eichenbaum, and Trabandt (2015) and Smets and Wouters (2007). Subsequent work by Gertler and Leahy (2008), Midrigan (2011), Alvarez and Lippi (2014) and others has painted a more nuanced picture, highlighting how the answer to the aforementioned question may depend on a number of hard-to-measure “details,” such as the stochastic properties of the idiosyncratic shocks to production costs or the possibility that firms may sell multiple goods and may face economies of scale in repricing all of them at once as opposed to each one separately. Nevertheless, it seems fair to say that the level of the nominal price rigidity required by quantitative versions of the New Keynesian framework is hard to reconcile with the available microeconomic evidence. The theoretical foundation and quantitative importance of nominal wage rigidity is another contentious issue in the literature.

But even if one takes for granted that large nominal rigidity exists at the macro level, this is not sufficient. The New Keynesian formalization of demand-driven fluctuations requires, not only monetary non-neutrality, but also sufficiently large countercyclical movements in the gap between the realized markup, $\mu_t$, and its flexible-price counterpart, or, equivalently, in the gaps of output, employment, and the real interest rate from their “natural-rate” counterparts.
Unfortunately, none of these gaps are directly observable. Any test of the empirical validity of the New Keynesian formalization of demand-driven fluctuations must therefore rely on strong assumptions about how these gaps can be proxied in the data. This is essentially the same difficulty as the one faced when trying to test the New Keynesian Philips Curve (NKPC). Let me elaborate.

Define \( x_t \equiv \log \mu^*_t - \log \mu_t \) as the (negative of the) gap between the realized markup and its flexible-price counterpart. In the literature, this is often referred to as the real marginal cost. In the infinite-horizon New Keynesian model, the NKPC takes the following form:

\[
\pi_t = \kappa x_t + \beta \mathbb{E}_t[\pi_{t+1}],
\]

or, by iterating,

\[
\pi_t = \kappa \sum_{k=0}^{\infty} \beta^k \mathbb{E}_t[x_{t+k}],
\]

where \( \pi_t \) denotes the inflation rate between period \( t - 1 \) and period \( t \), \( \mathbb{E}_t[\cdot] \) is the rational expectation operator, \( \kappa > 0 \) is a fixed scalar that parameterizes how responsive inflation is to innovations in the aforementioned gap, or the real marginal cost, and \( \beta > 0 \) is the discount factor. In the simplified, two-period version of the New Keynesian model used in this section, I have refrained from spelling out the price-setting behavior of the firms. I can nevertheless proxy for the NKPC by letting monetary policy replicate flexible prices in the second period, so that \( \mu_2 = \mu_2^* \) and \( x_2 = 0 \), and by imposing that the first-period inflation rate is given by \( \pi_1 = \kappa x_1 = \kappa (\log \mu^*_1 - \log \mu_1) \).

One way or another, the key observation is that, although the gap, \( x_t \), is not directly observable, this gap oughts to manifest in inflation. This has a simple interpretation. Fix a period and consider a firm that has the option to reset its price. The firm understands that it will likely be unable to adjust its price for a while. The firm also understand that, for any given price, a higher level of demand translates into a lower realized markup. To avoid such erosion in its realized markup, the firm sets a higher price when it expects a higher level of demand. As this logic applies to all the firms that have the option to reset, the price level today—and, equivalently, the inflation rate between yesterday and today—is an increasing function of the expected gaps over the relevant horizon.

This prediction is at the core of the New Keynesian model. If the bulk of the observed business cycles are driven by fluctuations in the gaps between sticky- and flexible-price allocations, it has to be that booms are periods of high inflation and recessions are periods of low inflation. What is more, the converse is also true: according to the model, positive [respectively, negative] innovations in inflation indicate positive [respectively, negative] innovations in the aforementioned gap.\(^{30}\)

\(^{30}\)To be precise, these statements are valid as long as the gaps are positively correlated over time: if positive gaps today are systematically followed by negative gaps tomorrow, inflation does not have to move with the current gap. I am ruling out this theoretical possibility because I can see no empirical justification for it. I am also ignoring variation in expectations of the central bank’s long-term inflation target, because I doubt that this is plausible at the business-cycle frequency.
Is this prediction borne by the data? As anticipated, answering this question is essentially the same as testing the validity of the NKPC—or of the older Philips curve, which abstracted from expectations, or of a number of variants that have appeared over the years in the literature. And the key challenge is that the gap $x_t$ is not directly unobservable.

Proxying $x_t$ with measures of the output gap published by the Fed reveals a major challenge: inflation appears to be negatively related with the gap, which is the exact opposite of what the theory requires. Gali and Gertler (1999) review this fact, argue that it is due to the poor quality of the considered empirical proxy, and proceed to offer alternative evidence in support of the NKPC. That evidence, however, relies on proxying $x_t$ with the labor share, an approach that rests on the untestable assumption that the equilibrium labor share would been constant if prices had been flexible. What is more, even if one takes for granted this assumption, the empirical relation between the labor share and inflation is rather weak. For instance, Angeletos, Collard, and Dellas (2017) use a Structural VAR approach to document that the macroeconomic time series can be described as follows: although more than 90% of the variation of the labor share at business-cycle frequencies can be accounted by a single shock, this shock explains less than 5% of the corresponding variation in inflation.

One can try to salvage the NKPC by arguing that most of the observed variation inflation is driven, not by variation in aggregate demand, but rather by the so-called cost-push shocks. To see what this means, rewrite the NKPC as follows:

$$\pi_t = \kappa \hat{x}_t + \beta E_t \pi_{t+1} + u_t,$$

where $\hat{x}_t$ is an empirical proxy for $-\mu_t$ and where $u_t$ is the cost-push shock, defined in the baseline model as $\kappa \mu_t^e$. In more flexible interpretations of the NKPC, $u_t$ could also capture measurement error, variation in expectations of the central bank’s long-run inflation target, irrational mistakes in predicting future inflation, and any other deviation from the theory. The data can then be described as follows: the slope, $\kappa$, is almost zero and the residual, $u_t$, accounts for almost all of the variation in $\pi_t$.\footnote{\cite{Sbordone2002}}

It is hard, at least for me, to view any of the above as empirical vindication of the New Keynesian formalization of demand-driven fluctuations. One may counter-argue that I am taking the NKPC too seriously: perhaps the empirical failures of old and new Philips curves represent only an indication that we lack a good theory of inflation, not an indication that we lack a good theory of the business cycle.

\footnote{A closely related approach is taken in Sbordone (2002).
Related, King and Watson (2012) argue that the findings of Gali and Gertler (1999) are driven by the fact that current inflation, $\pi_t$, is strongly correlated with the econometrician’s estimate of future inflation, $E_t \pi_{t+1}$, which by itself speaks very little to the core question of whether inflation predicts the labor share (the proposed measure of the gap).
In line with this observation, the extant DSGE literature (e.g., Smets and Wouters, 2007) attributes almost the entirety of the inflation fluctuations to markup shocks. See also the review of the empirical literature in Mavroeidis, Plagborg-Møller, and Stock (2014), as well as the topical discussion in Blanchard, Cerutti, and Summers (2015). Blanchard (2017) tries to salvage the natural-rate hypothesis as a gauge for monetary policy from the failures of Philips curves, but I find it hard to understand the one without the other.}
cycle. However, because the New Keynesian model equates demand-driven business cycles to gaps, which in turn are equated to inflation, it is quite delicate, if not incoherent, to claim that the model offers a satisfactory theory of demand-driven business cycles when the NKPC fails rather spectacularly.

The main challenge goes beyond these quibbles with the NKPC. When I look at the macroeconomic time series, I see strong co-movement between employment, output, consumption and investment, but little co-movement between these quantities and either productivity or inflation. This is illustrated in Figure 2, which takes the US time series and reports the scatterplots of the business-cycle component of output (on the horizontal axis) against the business-cycle components of hours worked, investment, consumption, TFP, labor productivity, and inflation (on the vertical axis). The top three panels reveal the strong co-movement of the real quantities; the bottom three panels reveal the absence of commensurate co-movement with TFP, labor productivity, and inflation. By this measure, the failure of the New Keynesian model is almost as profound as that of the RBC model.

Figure 2: Business-Cycle Comovement with Output (or lack thereof).

---

34The data are in quarterly frequency and cover the 1960-2015 period. Output is measured by GDP; hours worked by the hours of all persons in the non-farm business sector; consumption by the sum of personal consumption expenditures in nondurables goods and services; investment by the sum of personal consumption expenditures on durables goods, fixed private investment and changes in inventories; TFP by the utilization-adjustment measure provided in Fernald (2014); labor productivity by the ratio of GDP to total hours; and inflation by the change in the CPI index. The business-cycle components are obtained by applying the Band-Pass filter and isolating the frequencies corresponding to 6-32 quarters, as in Stock and Watson (1999). The results are nearly identical if the HP filter is applied instead.
To sum up, when I digest the lessons of the literature, look at Figure 2, or try to understand why the Great Recession was neither the Great TFP Loss nor the Great Deflation, I feel compelled to move beyond both the RBC model and the New Keynesian model. And while I recognize the evidence about monetary non-neutrality, as documented in the extant empirical literature or apparent in concrete episodes such as the Volker disinflation, I am not convinced that nominal rigidity is the most essential, and certainly not the only, reason for why demand shocks can trigger business cycles. Instead, like Beaudry and Portier (2013), I believe that the data demand a theory of non-inflationary, non-monetary, demand-driven fluctuations. I review how my research offers such a theory in the next section.35

3  Shifting the Focus from Nominal Rigidity to Frictional Coordination

The theory that I discuss in this section has two distinctive features. First, it reconciles the Keynesian narrative with flexible prices and, as result, allows it to be disconnected from the observed movements in inflation. Second, it gives a central position to the idea that the coordination of beliefs and behavior attained in the real world may be far less perfect than the one assumed in workhorse models of either the RBC model or the New Keynesian type.36

I will elaborate on what I mean by the latter statement in a moment. But let me first highlight the connection to, and the difference from, an earlier literature that sought to capture the role of coordination in models featuring multiple equilibria and sunspot fluctuations.37 In this literature, a coordination failure was said to obtain when one equilibrium was played rather than a “better” one; and demand-driven business cycles were triggered by sunspots that caused shifts in equilibrium outcomes without any shift in the underlying fundamentals such as preferences and technologies.

This literature was quite live in the 80’s and early 90’s, as part of early attempts to salvage the Keynesian narrative in the aftermath of the rational-expectations/RBC revolution. It nevertheless went out of fashion over time, because of a number of reasons, including: the sensitivity of policy predictions on seemingly arbitrary equilibrium selections; the delicate question of how one could conduct quantitative analysis without knowing how to choose among the many equilibria; the lack of solid empirical foundations; and the emergence and eventual dominance of the New Keynesian framework.

35Throughout this subsection, I have hammered the New Keynesian framework, and especially the NKPC, in order to motivate the alternative theory reviewed in the next section. My work, however, also offers two modifications of the NKPC that help improve its empirical performance. The one boils down to reducing the responsiveness of inflation to news about real marginal costs and output gaps (Angeletos and Liang, 2016a; Angeletos and Huo, 2017). The other develops a microfoundation of cost-push shocks in terms of correlated mistakes in expectations (Angeletos and La’O, 2009; Angeletos and Huo, 2017). See also the discussion in Section 4 for additional ways in which frictions in information and coordination can help improve the overall empirical performance of the New Keynesian framework.

36These points apply regardless of whether one abstracts from financial frictions, as I do here for simplicity, or incorporates them, as it would be necessary to do in order to talk meaningfully about the Great Recession.

37See, inter alia, Diamond (1982); Cass and Shell (1983); Cooper and John (1988); Guesnerie and Woodford (1993); Woodford (1991); Benhabib and Farmer (1994, 1999).
The latter shifted the focus away from coordination failure to nominal rigidity.

My research during the last few years is devoted on shifting the focus back to coordination failure. However, instead of equating coordination failure to equilibrium multiplicity, I equate it to lack of common knowledge and higher-order uncertainty within unique-equilibrium models.

Both the RBC framework and the New Keynesian framework—whether in the simplified forms I described earlier on or the various richer forms found in the literature—impose that all agents in the economy share the same information at all points of time. Together with rational expectations, this implies that all agents can reach a "perfect consensus", not only about the exogenous shocks hitting the economy, but also about the endogenous state of the economy in the present and its likely path in the future. In this sense, both the RBC and the New Keynesian framework impose that the agents can perfectly coordinate their beliefs and actions along the equilibrium.

By contrast, my research allows this coordination to be imperfect by letting the agents to have heterogeneous information about the current state and/or the future prospects of the economy and, thereby, to face uncertainty about one another's beliefs and actions. This approach, which builds heavily on Morris and Shin (1998, 2002, 2003) and Woodford (2003a), permits one to accommodate frictional coordination within workhorse macroeconomic models. It can be viewed as more robust than the older approach that rested on multiple equilibria, because appropriate perturbations of the information structure, of the type considered in the global-games literature, can transform any model to a model with a unique equilibrium (Weinstein and Yildiz, 2007).\footnote{As discussed next, however, that the essence of multiple-equilibrium models is preserved even when global-games techniques are used to select a unique equilibrium, because the unique equilibrium can vary with shocks that resemble animal spirits. See also the complementary discussions in Morris and Shin (2002) and Angeletos and Werning (2006) about the sunspot-like function of public information in environments in which coordination is important.} It helps reveal how crucially the predictions of standard workhorse macroeconomic models depend on the conventional but unrealistic assumption that all the agents have a homogenous understanding of the current state and the future prospects of the economy. And it leads to a parsimonious explanation of multiple empirical regularities as well as to new policy insights.

### 3.1 Beauty Contests and Sentiments

I now discuss how my approach opens to door to a certain kind of fluctuations that resemble the product of animal spirits and that help reconcile the Keynesian narrative with the RBC framework. This discussion is based on a stripped down version of my work with Jennifer La’O on “sentiments” (Angeletos and La’O, 2013).

The economy has a large number of agents, who can be thought of as both consumers and producers; let’s call them “farmers.” Each farmer produces a single good, using his own labor, but wishes to consume also the good of another, randomly selected, farmer. The farmers therefore engage in barter
trade through random pairwise matching. The terms of trade within each match are determined in a competitive fashion. These assumptions are deliberately unrealistic: they seek to keep the analysis as close as possible to those found in textbook treatments of the Edgeworth box and of demand and supply. The key novelty is that any two farmers are allowed to have differential information about the underlying state of Nature and, as result, can face higher-order uncertainty about their likely terms of trade—or, equivalently, about demand and supply.

Suppose, in particular, that each period can be split into two sub-periods, the “morning” and the “afternoon.” Production takes place in the morning; trade and consumption take place in the afternoon. Importantly, each farmer decides how much effort to exert and how much to produce prior to observing his exact match and the terms of trade; in this sense, supply is determined under incomplete information about demand.

Because of space constraints, I skip the details of the underlying micro-foundations. For the present purposes, it suffices to note that the equilibrium of the considered model boils down to the solution of the following fixed-point problem:

\[ y_{it} = \varphi A_{it} + \alpha E_{it}[y_{jt}], \]  

(12)

where \( y_{it} \) is the output of farmer \( i \) in period \( t \), \( A_{it} \) is her productivity in that period, \( E_{it} \) is her rational expectation in the morning of that period, \( j \) stands for the identity of her random trading partner, \( y_{jt} \) is the output of her trading partner, and \( \varphi > 0 \) and \( \alpha \in (0,1) \) are scalars that depend on deeper preferences and technology parameters and that can be treated as exogenous parameters in the present discussion.

Condition (12) states that the equilibrium output of a farmer is an increasing function of her productivity and of her expectation of the output of her trading partner. Why? Because higher productivity means a lower cost of producing and because a higher level of production by her trading partner translates to higher demand for her own product (or, equivalently, better terms of trade).

This condition can be thought as the best response condition in a two-player game: the players are the farmers within a match and their actions are the level of production. This game features strategic complementarity and linear best responses. It is therefore closely related to the class of “beauty contests” studied in Morris and Shin (2002), Angeletos and Pavan (2007, 2009), and Bergemann and Morris (2013). Here, strategic complementarity emerges simply because higher supply from one farmer translates to higher demand for another farmer.

Because \( \alpha \in (0,1) \), condition (12) defines a contraction mapping. The equilibrium outcome is unique, it coincides with the unique rationalizable outcome, and it is pinned down by the hierarchy of beliefs about the underlying fundamentals (the farmer-specific productivities and the realized matches). To see this more clearly, suppose that any two farmers \( i \) and \( j \) that have been matched to-
gether have common knowledge of their identities but not necessarily of their productivities. Then, by iterating (12), we readily see that \( i \)'s output is given by

\[
y_{it} = \phi A_{it} + \phi \alpha E_{it} [A_{jt}] + \phi \alpha^2 E_{it} [E_{jt} [A_{it}]] + \phi \alpha^3 E_{it} [E_{jt} [E_{it} [A_{jt}]]] + \ldots
\]

In short, \( i \)'s output depends, not only on her own productivity and on her belief of the productivity of her trading partner, but also on her belief of what her partner may himself think about his likely match.

The above is true regardless of the information structure. The information structure, however, determines whether coordination is “perfect” or “imperfect” in the following sense. When information is complete (i.e., all farmers share the same information about their matches, about one another's productivities, and the entire state of Nature), all the higher-order beliefs collapse to the true fundamentals. As a result, the farmers face no uncertainty about one another's choices and aggregate output is pinned by fundamentals (TFP), as in the standard RBC model. By contrast, when information is incomplete, the farmers face uncertainty about one another's beliefs and choices. This uncertainty formalizes the precise sense in which coordination is imperfect. It also rationalizes correlated “mistakes” in the forecasts that farmers make when trying to predict one another's choices. These mistakes in turn manifest as a type of fluctuations.

To see this more clearly, let the realized productivities \((A_{it}, A_{jt})\) be common knowledge within each match \((i, j)\). Then, condition (13) holds with

\[
E_{it} [A_{jt}] = A_{jt}, \quad E_{it} [E_{jt} [A_{it}]] = A_{it}, \quad E_{it} [E_{jt} [E_{it} [A_{jt}]]] = A_{jt},
\]

and so on. But then we have

\[
y_{it} = \phi \left\{ \sum_{h=0}^{\infty} \alpha^{2h} A_{it} + \sum_{h=0}^{\infty} \alpha^{2h+1} A_{jt} \right\} = \frac{\phi}{1 - \alpha^2} \left\{ A_{it} + \alpha A_{jt} \right\}
\]

and similarly \( y_{jt} = \frac{\phi}{1 - \alpha^2} \left\{ A_{jt} + \alpha A_{it} \right\} \). And since this is true for every match \((i, j)\), aggregate output is given by

\[
y_t = \frac{\phi}{1 - \alpha} A_t,
\]

where \( A_t \) denotes aggregate TFP. In short, the complete-information version of the consider model is a close cousin for the RBC model: it attributes the business cycle to the familiar kind of “supply shocks.”

When, instead, the farmers lack common knowledge of each other's productivities, higher-order beliefs can differ from first-order beliefs. What is more, the variation in higher-order beliefs need not be

\[39\text{This requires the productivities themselves to be random.}\]
spanned by the variation in either the underlying fundamentals or the first-order beliefs: higher-order beliefs can vary for seemingly extrinsic reasons. To illustrate, suppose that every farmer $i$ observes two noisy signals about its likely partner $j = m(i, t)$. The first signal is given by

$$x_{it}^1 = A_{jt} + \epsilon_{it},$$

and can be thought of as a signal of the trading partner’s productivity. The second signal is given by

$$x_{it}^2 = \epsilon_{jt} + \xi_{it},$$

and can be thought of as a signal of the error in trading partner’s signal (or, equivalently, of the associated “mistake” in her choices). Suppose further that $\xi_{it}$, the error in the second signal, is correlated across all the farmer in the economy. For instance, suppose that $\xi_{it} = \xi_t$, where $\xi_t$ is an aggregate shock. Suppose further that the latter shock is uncorrelated with aggregate productivity. It follows that $\xi_t$ represents an independent shock to higher-order beliefs: when the realized $\xi_t$ is higher, the first-order beliefs $E_{it}[A_{jt}]$ and $E_{jt}[A_{it}]$ do not move, but the second-order beliefs $E_{it}[E_{jt}[A_{it}]]$ and $E_{jt}[E_{it}[A_{jt}]]$ go up, and so do the belief of third and higher orders.

Angeletos and La’O (2013) refer to $\xi_t$ as a “sentiment shock” because, in equilibrium, this shock helps capture the sentiment (belief) that a farmer has about her terms of trade and the returns to her production. In particular, a positive $\xi_t$ realization captures states of Nature in which the average farmer overproduces relative to the frictionless RBC benchmark because, and only because, she is optimistic the other farmers also overproduce. And symmetrically, a negative $\xi_t$ realization captures states of Nature in which the average farmer cuts down her production because, and only because, she is pessimistic that the other farmers will act similarly. In this sense, it is as if the economy fluctuates in response to “animal spirits.” Furthermore, because these fluctuations are associated with variation in the expected and the realized demand faced by the average farmer for given supply conditions (i.e., for given technology and marginal costs), these fluctuations can be said to have a Keynesian flavor.

To recap, the kind of “sentiment shocks” introduced in Angeletos, Collard, and Dellas (2015) help capture waves of optimism and pessimism that are disconnected from fundamentals and that resemble animal spirits, but are consistent with equilibrium uniqueness. Benhabib, Wang, and Wen (2015) develop a more realistic micro-foundation of the information structure and push further the interpretation of this kind of shock as a form of “aggregate demand shock.” Finally, Angeletos, Collard, and Dellas (2015) and Huo and Takayama (2015) explore the observable implications and the quantitative potential of this kind of shock; I review some of the findings next.
3.2 Sentiments and the Business Cycle

In joint work with Fabrice Collard and Harris Dellas, I have developed a tractable method for augmenting a large class of linear DSGE models with rich dynamics in higher-order beliefs. The method leverages on a certain departure from the common prior and rational-expectations assumptions in order to maximize tractability and ease the simulation and the structural estimation of both small and large models. In “Quantifying Confidence” (Angeletos, Collard, and Dellas, 2015), we use this method to shed light on the observable implications and the quantitative potential of extrinsic shocks to higher-order beliefs, of the kind described above.40

Let me review some of the findings. Start with the baseline RBC model, which attributes the entirety of the business cycle to aggregate TFP shocks. Modify this model by removing common knowledge of the realized TFP shock and by allowing for an aggregate shock to higher-order beliefs. By the latter I mean a shock that is orthogonal to the TFP shock, does not affect the (first-order) beliefs that the agents form about TFP, and nevertheless triggers transitory variation in the beliefs that the agents form about the beliefs of others. This shock is therefore analogous to the sentiment shock formalized in Angeletos and La’O (2013) and reviewed in the previous subsection. The exact modeling is different—instead of pairwise matches among farmers, there is the standard interaction of multiple households and firms in labor and capital markets—but the essence is the same: the considered shock helps accommodate extrinsic waves of optimism and pessimism about the choices of others and, thereby, about aggregate economic outcomes.

What are the observable implications of this kind of shock? Figure 3 addresses this question by reporting the impulse response functions of the model’s key endogenous outcomes to a negative sentiment shock. Output, consumption, investment, and employment go down, while TFP remains stable and labor productivity slightly increases. This co-movement in key macroeconomic quantities without commensurate co-movement in TFP or labor productivity matches our “intuitive” notion of an adverse demand shock, as well as the empirical regularities that are associated with that notion (e.g., Blanchard and Quah, 1989).

The mechanism is the following. By construction, the shock causes higher-order beliefs of TFP to fall. In equilibrium, this triggers a wave of pessimism about the short-run economic outlook without any commensurate movement in either the actual TFP and other fundamentals or beliefs thereof. In particular, the firms become pessimistic about profitability and returns over the next few quarters, and the consumers to become pessimistic about employment and income over the same horizon, without any change in their expectations of either the exogenous fundamentals or of the endogenous outcomes in the medium to long run.

40The method, however, is more general and can be used for other purposes too. For instance, it can help quantify the sluggish response of beliefs to intrinsic shocks, as in Woodford (2003a), Nimark (2008, 2017), and Angeletos and Lian (2016a). It can also help connect the theory with survey evidence on expectations.
How does such a wave of pessimism affect actual outcomes? Because firms expect the demand for their products to be relatively weak over the next few quarters, they find it optimal to lower their own demand for labor and capital. As a consequence, households expect to experience a transitory fall in wages, capital returns, and overall income. Because this entails relatively weak wealth effects and relatively strong substitution effects, households react by working less and by reducing both consumption and saving.

The belief waves described above are therefore able to generate the empirical patterns seen in Figure 2: strong positive co-movement between employment, output, consumption, and investment at the business-cycle frequency, without commensurate movements in labor productivity, TFP, and inflation at any frequency.

So far, I have shown how to accommodate a Keynesian type of fluctuations within the RBC model, while abstracting from nominal rigidities. It is straightforward to add Calvo-like sticky prices; this gives the baseline New Keynesian model augmented with sentiment shocks. When monetary policy replicates flexible prices, one gets (trivially) the same response in the real quantities along with no response in inflation. This explains how sentiment shocks can serve as non-inflationary demand shocks within the New Keynesian model and can therefore help address some of its empirical shortcomings.

What is more, one can show that the real effects of a sentiment shock under flexible prices are similar to those of a monetary shock under sticky prices. This underscores how the mechanism described above is a close cousin to the one at the core of the New Keynesian framework—except for the fact that it does not rest on nominal rigidity, policy constraints, and commensurate inflation movements.

To connect this point to the discussion of Section 2, consider the baseline model in Angeletos, Collard, and Dellas (2015). This is essentially the multiple-period version of the RBC model introduced in Subsection 2.2, except for the replacement of inter-temporal preference shocks with sentiments shocks. The equilibrium quantities can be shown to satisfy the following conditions, for all $t$:

$$
\frac{v'(\ell_t)}{u'(c_t)} = (1 - \tau_t^f) A_t F_n(k_t, n_t) \quad \text{and} \quad u'(c_t) = \beta E_t \left\{ u'(c_{t+1}) \left[ 1 + (1 - \tau_t^f) A_{t+1} F_k(k_{t+1}, n_{t+1}) \right] \right\},
$$
where the τ’s capture the endogenous wedges induced by the sentiment shock. In particular, a negative sentiment shock manifests as a joint increase in τ_t^f and τ_t^k: pessimistic beliefs about the choices of others are akin to a joint tax on labor and capital.41

These belief-induced wedges play a similar as the realized monopoly markup in the New Keynesian model: they encapsulate the output gaps that obtain relative to the predictions of the baseline RBC model. But whereas in the New Keynesian model the wedges and the gaps are the symptom of nominal rigidity, in our model they are the symptom of a friction in the coordination of the beliefs and the economic decisions of a diverse population. And whereas in the New Keynesian model the wedges and the gaps ought to manifest in inflation (through the NKPC), in our model they do not.

To elaborate on the quantitative content of all these observations, Angeletos, Collard, and Dellas (2015) consider a horserace between the version of the RBC model that contains the sentiment shock with versions of the New Keynesian model that rule out the sentiment shock and, instead, feature more standard formalizations of “demand shocks,” such as the kind of discount-factor shock discussed earlier on. For comparable calibrations, the RBC model with the sentiment shock outperforms its New Keynesian competitors in terms of matching the key business-cycle moments in the data. This is, not only because the former model does not have to rely on counterfactually large movements in inflation, but also because the sentiment shock is better able to generate the co-movement patterns among the real quantities seen in the data. For essentially the same reason, the sentiment shock emerges as the main driver of the business cycle in estimated, medium-scale models that allow for a multitude of other shocks and for some of the familiar bells and whistles of the DSGE literature.

These findings raise the following question. Is the business cycle “truly” driven by the considered kind of sentiment shock, or is the latter a proxy for something else? At this point, I can not provide a convincing answer to this question. Part of the problem is that the sentiment shock, like any other structural shock, is not directly observed by the econometrician. That said, the combination of the quantitative findings reported above, the complementary findings in Huo and Takayama (2015), and the empirical evidence reported in Angeletos, Collard, and Dellas (2017), Beaudry and Portier (2013), and Beaudry, Galizia, and Portier (2015) seem to point in the following conclusion. There are important empirical regularities for which the state of the art has failed to provide a parsimonious structural explanation. The theoretical mechanism I have proposed does not have to be the full story, but it certainly helps fill the void.42

41This follows from Subsection 5.4 of Angeletos, Collard, and Dellas (2015), which discusses how sentiments shocks manifest as wedges in terms of business-cycle accounting (Chari, Kehoe, and McGrattan, 2007).

42More direct evidence in favor of, or against, the proposed mechanism could potentially be provided by identifying, through VARs or other “atheoretical” methods, waves of optimisms and pessimisms about the short-run economic outlook that are orthogonal to technology, monetary policy, and other fundamentals. The inherent difficulty with this approach is that the very definition of what is a “fundamental” and what is an “outcome” depends on the class of models one has in mind. That said, the combination of the macroeconomic time series with survey evidence on expectations could help shed further light on the determinants and the consequences of expectations.
3.3 From Shocks to Propagation

So far I have tried to make sense of demand-driven business cycles using an extrinsic sentiment shock. In general, a shock in a model is proxy for a force, or propagation mechanism, whose deeper micro-foundations the theorist abstracts from in order to make progress in understanding its consequences. In the context of interest, the sentiment shock maybe a crude proxy for waves of optimism and pessimism caused by more familiar triggers, such as a shock to consumer credit.

Chen Lian and I have been exploring this idea in ongoing research (Angeletos and Lian, 2016c). We argue that the combination of frictional coordination with the confusion of idiosyncratic and aggregate shocks can generate realistic business cycles out of the kind of consumer-spending (or discount-rate) shocks described before. We further show how the same ingredients give rise to a feedback mechanism that resembles the Keynesian multiplier, despite the absence of any kind of nominal rigidity.

The model considered in Angeletos and Lian (2016c) has the same neoclassical backbone as the model studied in Subsection 2.2, except that there are now a large number of consumers and producers that interact in decentralized markets (“islands”). Each household’s preferences take the same form as in condition (1) and are subject to a discount-rate shock. The latter has both an aggregate component (proxying for aggregate shocks to consumer credit and aggregate demand) and an idiosyncratic component (proxying for, say, borrowing heterogeneity). Each household also contains a single worker-farmer, who produces only one of the many varieties that are consumed by every household. Finally, there are good-specific demand and supply shocks, in the form of, respectively, an economy-wide but good-specific taste shock and a farmer-specific productivity shock.

Since the aggregate discount-rate shock is the only aggregate shock in the economy, the entire variation in aggregate quantities has to be driven by it. But whether a given shock causes a boom or a recession depends critically on the whether information is complete or not.

When information is complete, the model reduces, in effect, to the one studied in Subsection 2.2. This leaves no room for the Keynesian narrative: for the reasons already explained, the drop in consumer spending comes together with a boom in aggregate employment, investment, and output.

This changes when information is incomplete. The discount-rate shock triggers a drop in the demand faced by each farmer (or firm). But because information is incomplete, some farmers may not be sure whether the drop in their demand is due to idiosyncratic or aggregate reasons. As a result, some farmers may work less and may also instruct their sibling-consumers to spend less. But as these consumers spend less, other farmers experience a further drop in their demand. These farmers may now find it optimal to work less and to instruct their own siblings to spend less, even if they themselves are fully aware that the initial trigger was an aggregate discount-rate shock. An extra round of reduction in output, labor, and consumption therefore takes place. And so on.

A more realistic version of the theory replaces the farmers with collections of firms and workers...
(and adds labor markets). In response to the aggregate discount-rate shock, some firms see a decrease in the demand for their products and start hiring less. Some workers see wages go down, or unemployment go up, and start spending less. Additional firms then see their demand go down and respond by contributing to even less hiring. And so on. This explanation abstracts from investment; the logic, however, directly extends from the firms’ incentives to hire to the firms’ incentives to invest.

The mechanism described above has a sharp Keynesian flavor. One may even argue that our model offers a more “faithful” formalization of the considered narrative than the New Keynesian model: the mechanism draws directly from the elementary demand-and-supply reasoning that is familiar from microeconomics. But how exactly is such PE reasoning reconciled with the GE response of the economy? In other words, why is this response different from the one characterized in Subsection 2.2?

Part of the answer rests on the assumed inability of the firms (or the farmers) to tell apart the sources of variation in their demand and the similar inability of the consumers to tell apart the sources of variation in their income. This ingredient of our theory is similar to Lucas (1972), except that the agents in our model are confusing different kinds of real terms as opposed to confusing nominal terms for real terms. What is more, this confusion does not have to the product of segmented market interactions and missing public signals; it can be the product of rational inattention or, even more erratically, the product of bounded rationality.

But this is not the whole story; it is only the starting point. The most novel and, in our view, the most intriguing part of the theory has to do with the feedback loops described above. Because these feedback loops are akin to strategic complementarity in “beauty contests,” the confusion of some agents rationalize a similar behavior by other agents regardless of whether the latter are also confused or not. It is this part of our theory that formalizes the Keynesian multiplier.

As a matter of fact, the mechanism is valid even if no agent is actually confused, provided that this fact is not common knowledge. In this sense, the key is not the Lucas-like confusion per se, but rather the inability of the agents to coordinate on the kind of GE response predicted by the standard RBC framework. In short, it is as if the GE forces of that framework have been attenuated and, instead, the PE logic of the Keynesian narrative has prevailed.

4 Expectations and GE Adjustment

I now elaborate on the broader idea that lack of common knowledge attenuates GE effects by arresting the adjustment of expectations to aggregate shocks. I first articulate the basic idea. I then discuss how this idea sheds new light on the question of how monetary and fiscal policy influence aggregate demand. I finally discuss how a variant of this idea offers an empirically plausible micro-foundation of some of the “bells and whistles” of the New Keynesian model.
4.1 Dampening GE

In prior work with Chen Lian (Angeletos and Lian, 2017), I have formulated the basic idea that lack of common knowledge attenuates GE within the context of an elementary Walrasian economy, featuring decentralized and sequential trading. There are two periods, “today” and “tomorrow.” There are three goods, which is the minimum necessary for distinguishing PE from GE. (Henceforth, GE and PE as acronyms for, respectively, general and partial equilibrium.) One of the goods serves as the numeraire and can be consumed in both periods; think of it as leisure. The other two goods are specific to two periods; think of them as “today’s goods” and “tomorrow’s goods.” Finally, there is a large number of “marketplaces,” and every agent can trade in a single marketplace in each period but may randomly move from one marketplace to another as time passes.43

How does this economy respond to an aggregate demand shock, namely, a shock that shifts the local demand for today’s goods in all (or many) marketplaces? We first address this question in a “frictionless” benchmark that exemplifies the modeling practice in the majority of applied work. This benchmark is defined by imposing Rational Expectation Equilibrium together with common knowledge of the aggregate shock. Its predictions are illustrated in Figure 4.44

Consider an arbitrary marketplace $m$ during the first period. We denote the local price of today’s goods by $p_m$ and the local quantity by $q_m$. The pre-shock demand and supply curves are illustrated by the solid lines in the figure. Had the shock being idiosyncratic (specific to marketplace $m$), it would have shifted only the local demand curve and it would have only a PE effect. This effect is represented by the movement of the market-clearing pair $(p_m, q_m)$ from point $X$ to point $Y$ in either panel of the figure. Because we are considering an aggregate shock, however, there is an additional GE effect, which has to do with the concurrent adjustment in aggregate economic outcomes and in the prices that agents expect to face tomorrow. This effect triggers a shift in the supply curve, as well as a further shift in the demand curve; it is represented by the movement from point $Y$ to point $Z$. In the left panel, this kind of GE adjustment amplifies the PE effect; in the right panel, it mitigates it.

The points made so far should be familiar: GE mechanisms can either amplify or offset PE effects. A novel lesson emerges once we relax the assumption that the shock is common knowledge.

This lesson can be summarized as follows. The rational expectations hypothesis alone does not nail down the relevant GE effect. It only restricts its (absolute) magnitude within an interval. By imposing rational expectations together with common knowledge of the shock, standard modeling practices pick, perhaps inadvertently, the upper bound of this interval. We instead show how one can span the entire interval by removing common knowledge of the shock. In terms of the figure, this means that

---

43These assumptions are deliberately stark. They nevertheless capture two basic facts: that most trading is decentralized; and that agents care, but are uncertain, about the behavior of agents they do not currently trade with. They also help us draw a clear line between partial and general equilibrium: the former refers to the adjustment of a single marketplace in isolation of the rest of the economy, the latter to the joint adjustment of all the marketplaces.

44Relative to the previous parts of this article, I am now switching from a negative to a positive shock.
the GE adjustment of the (modified) economy can lie anywhere along the interval between $Y$ and $Z$.

The basic insight behind this result is the following. Regardless of the information structure, the rational-expectations hypothesis imposes a fixed-point relation between subjective beliefs and actual outcomes. But once agents lack common knowledge of the innovations in the underlying fundamentals, this fixed point is pinned down, not only by what the agents know about these innovations, but also by what they think that others know, and so on. As one varies the degree of such higher-order knowledge, one also varies the potency of the relevant GE effect.

This explains the sense in which standard practice has “overstated” the importance of GE mechanisms and how relaxing common knowledge helps merge the gap between micro intuitions and macro predictions. Importantly, this is true regardless of whether the GE effects of a shock amplify its PE effects (left panel in the figure) or the opposite (right panel). I invite the reader to take a look at Angeletos and Lian (2017) for more details, for certain connections to empirical work, and for the comparison to recent work that touches on similar ideas by replacing the rational-expectations solution concept with Level-k Thinking (Garcia-Schmidt and Woodford, 2015; Farhi and Werning, 2017) or a certain form of myopia (Gabaix, 2016).

### 4.2 Application: Forward Guidance and Fiscal Stimuli

I now discuss how the aforementioned ideas shed new light on the question of how policy can influence aggregate demand. Consider the New Keynesian framework and ask how monetary policy influences economic activity within that framework. By changing the interest rate(s) faced by the typical consumer, monetary policy has a direct effect on the budget, the incentives, and the behavior of that agent. This kind of PE (or, more precisely, decision-theoretic) effect is relatively modest: it explains only a small part of the real effects of monetary policy. The largest part has to do with GE mechanisms, which act as multipliers of the underlying PE effect.
The most crucial among these GE mechanisms is the feedback loop between aggregate spending and inflation: reducing interest rates stimulates spending, which in turn raises inflation, which in turn reduces real rate further and stimulates spending even more, and so on. In the baseline New Keynesian model, this mechanism is captured by the interaction of the representative household’s Euler condition (the modern analogue of the IS curve) with the NKPC (the modern analogue of the Philips curve). But there are two additional GE mechanisms, buried underneath these equations. The one has to do with the feedback from future inflation to current inflation: for given real marginal costs, the individual firm is more willing to raise its nominal price today if she expects other firms to do the same in the future. The other has to do with the feedback from aggregate spending to individual spending: when the individual consumer expects other consumers to spend more, she is encouraged to spend more herself, because her own income increases with aggregate consumption.

Building on the more abstract insight described in the previous subsection, Angeletos and Lian (2016a) show how relaxing the common-knowledge assumption attenuates all the aforementioned GE effects. What is more, it is shown that the attenuation increases with the horizon that the agents have to forecast: it is as if the economy is populated by a representative agent that is myopic in the sense that she discounts the impact of the policies, or shocks, on future outcomes more heavily than what it is rational. This is because policies, or shocks, that work in long horizons map to beliefs of high order and, in this sense, require an implausible degree of coordination.

This offers a resolution to the so-called forward-guidance puzzle. The latter refers to the paradoxical prediction that, when the economy is at the Zero Lower Bound, a promise to keep interest rates low in the far future can have humongous effects on current economic activity. This prediction is at odds with the PE intuition that interest rates further into the future matter less because of discounting. It is thus driven exclusively by the GE effects described above and, in particular, by the property that these GE effects pile up as the horizon increases. Our findings reduce the gap between the predictions of the model and the more appealing PE intuition.

Our findings also offer a rationale for the front-loading of fiscal policy. The baseline New Keynesian model predicts that fiscal stimuli should be back-loaded in order to pile up the feedback loops between inflation and spending. By contrast, our modification allows for the possibility that fiscal stimuli should be front-loaded in order to minimize the bite of the coordination friction.45

More broadly, these findings indicate how the considered friction can slow down the adjustment of both aggregate demand and aggregate supply to shocks and policy shifts. This helps understand why the macroeconomic time series may feature more sluggish responses to aggregate shocks relative to those seen in the response of microeconomic outcomes to idiosyncratic shocks.46

45 This complements a different, more familiar, argument for the front-loading of fiscal stimuli, the one based on liquidity constraints and “hand-to-mouth” consumers (e.g., Gali, López-Salido, and Vallés, 2007).

46 This point also underscores a key difference between my approach and alternative approaches which, instead, focus on generating sluggishness at the microeconomic level, such as the literature on adjustment costs, the decision-theoretic
4.3 Anchored Expectations

The preceding discussion illustrates how relaxing common knowledge, and accommodating imperfect coordination, can modify the predictions of the New Keynesian framework in manners that appear to be both conceptually appealing and empirically plausible. Reinforcing this point, ongoing work with Zhen Huo (Angeletos and Huo, 2017) shows how incomplete information can offer a microfoundation for some of the more dubious “bells and whistles” that the New Keynesian framework requires in order to match the macroeconomic time series.

Consider, in particular, the specific kinds of adjustment costs in investment, consumption habit, and price indexation considered in Christiano, Eichenbaum, and Evans (2005), Smets and Wouters (2007), and almost all the subsequent DSGE literature. These modeling ingredients are essential for the ability of the New Keynesian framework to capture the response of the economy to identified monetary shocks, to generate realistic comovement in real quantities out of demand shocks, and, more broadly, to capture various salient features of the macroeconomic time series. And yet, these modeling ingredients have weak empirical foundations.

My work with Zhen Huo offers a potential resolution: incomplete information can be observationally equivalent to such bells and whistles vis-a-vis the aggregate time series.\(^{47}\)

To illustrate, consider the NKPC. As noted earlier, the standard version of the NKPC is given by

\[
\pi_t = \kappa x_t + \beta \mathbb{E}_t [\pi_{t+1}],
\]

(14)

where \(\pi_t\) denotes inflation, \(x_t\) denotes the output gap (aka the real marginal cost), \(\kappa > 0\) parameterizes the responsiveness of inflation to innovations in the gap, and \(\beta \in (0, 1)\) is the subjective discount factor. This condition follows from aggregating the optimal price-setting decisions of the firms and hinges on the forward-looking nature of these decisions. In particular, because the firms that have the option to reset their prices today understand that they are likely to be stuck with the same price for a while, they set their prices in proportion to their expectation of the discounted present value of the real marginal costs that they are likely to face in the future. It then follows that \(\pi_t\) depends, not only on the current value of \(x_t\), but also on the firms’ expectations of its future path, which in turn explains why the forward-looking nature of the NKPC.

These points are well understood. What is not well understood, however, is how the version of the NKPC given in condition (14) depends on the assumption that the firms have common knowledge of

\(^{47}\)The subsequent discussion focuses on the implications of adding informational frictions in the New Keynesian framework while maintaining the usual formalization of nominal rigidity, namely, Calvo-like sticky prices. The works of Woodford (2003a), Mankiw and Reis (2002), and Mackowiak and Wiederholt (2009) make a complementary but different contribution by showing how informational frictions may substitute for Calvo-like sticky prices as sources of nominal rigidity. Also note that these works abstract from the forward-looking aspect in the firms’ price-setting behavior. By contrast, such forward-looking behavior is central to the results reviewed here.
the current value \( x_t \) and a common belief about its future path. Without this assumption, the optimal price-setting behavior of each firm is still driven by her expectations of the discounted present value of her real marginal costs, which in turn implies that inflation is still driven by the average of these expectations in the cross section of firms, but this average expectation no more coincides with the expectation of a representative agent. As a result, condition (14) has to be modified.

Suppose, in particular, that \( x_t \) follows an AR(1) process, or a random walk, and that the information of every firm can be represented by a series of Gaussian private signals about \( x_t \). My work with Zhen Huo establishes that, under this scenario, the appropriate modification of the NKPC is as follows:

\[
\pi_t = \kappa' x_t + \beta' \mathbb{E}_t [\pi_{t+1}] + \gamma \pi_{t-1},
\]

where the scalars \((\kappa', \beta', \gamma)\) are pinned down by the underlying parameters \((\kappa, \beta)\) and the information structure. These scalars satisfy \( \beta' < \beta \) and \( \gamma > 0 \). It is therefore as if the firms discount the future more heavily \( (\beta' < \beta) \) and, in addition, condition their current price increases on the past inflation \( (\gamma > 0) \).

The first feature is for the reasons explained earlier: by arresting the adjustment of the expectations that the firms form about the behavior of other firms and the resulting inflation (a GE mechanism), the informational friction causes the economy to behave as if the firms were myopic. The second feature is due to the interaction of learning with higher-order uncertainty: as time passes and firms accumulate more information about the state of the economy and about one another’s responses, beliefs adjust only slowly towards their frictionless counterpart. It is therefore as if current beliefs and outcomes are anchored to past outcomes.\(^{48}\)

These features—discounting of the future and anchoring to the past—induce the kind of empirical patterns that the DSGE literature have sought to a variety of bells and whistles, including the so-called hybrid version of the NKPC and the specific forms of habit in consumption and adjustment costs in investment popularized by Christiano, Eichenbaum, and Trabandt (2015) and Smets and Wouters (2007). Of course, incomplete information can itself be viewed as yet another kind of bells and whistles. But whereas the DSGE literature requires multiple sets of bells and whistles, essentially a different one for each equation of the model, my work with Zhen Huo suggests that the same objectives can be accomplished with one friction. What is more, this friction seems consistent with the available evidence on the inertia of expectations, such as that documented in Coibion and Gorodnichenko (2012, 2015a) and Vellekoop and Wiederholt (2017). For additional work that indicates the value of introducing incomplete information and higher-order uncertainty in the New Keynesian framework, see Nimark (2008) and Wiederholt (2016).

\(^{48}\)Woodford (2003a) also emphasizes the inertia in the adjustment of beliefs, but does not identify the aforementioned discounting because he abstracts from forward-looking behavior. Conversely, Gabaix (2016) and Farhi and Werning (2017) consider two distinct departures from the rational-expectations hypothesis, both of which boil down to discounting future outcomes but, unlike the approach described here, do not generate the backward-looking element that the data call for.
5 Conclusion

In this article I surveyed, and advertised, my current research agenda. But I also tried to convey three broader lessons. First, augmenting workhorse macroeconomic models with incomplete information and higher-order uncertainty is a fruitful method for accommodating realistic frictions in coordination, for capturing salient features of the data, and for offering new guidance to policy. Second, the proposed approach helps reveal the “true” predictions of these models, in the sense of disentangling the rational-expectations hypothesis from the auxiliary assumption that the agents have common knowledge of the current state and future prospects of the economy. Finally, the familiar narrative that a drop in aggregate demand can cause a recession is far more subtle than what it appears to be at first glance. By addressing this narrative in both old and new ways, I hope to raise the general audience’s appreciation of its subtlety and to invite further research into its precise meaning.
References


Huo, Zhen and Naoki Takayama. 2015. “Higher Order Beliefs, Confidence, and Business Cycles.” mimeo, Yale University.


