FRICIONAL COORDINATION

George-Marios Angeletos
MIT

Abstract
The notion that business cycles are driven by demand shocks is subtle. I first review some of the conceptual and empirical challenges faced when trying to accommodate this notion in micro-founded, general-equilibrium models. I next review my own research, which sheds new light on the observed business cycles by accommodating frictional coordination in the form of higher-order uncertainty. This makes room for forces akin to animal spirits even when the equilibrium is unique. It allows demand shocks to generate realistic business cycles even when nominal rigidity is absent or undone by appropriate monetary policy. It modifies the general-equilibrium predictions of workhorse macroeconomic models in manners that seem both conceptually appealing and empirically relevant. And it offers new guidance to policy.

1. Introduction

What causes business cycles? Is it “supply shocks” such as those associated with changes in the general level of know-how or with mis-allocation in production (e.g., Chodorow-Reich 2014)? Or is it “demand shocks” such as those triggered by a drop in consumer sentiment (e.g., Lorenzoni 2009; Angeletos and La’O 2013) or by a collapse in housing wealth and debt deleveraging (e.g., Mian and Sufi 2014; Guerrieri and Lorenzoni 2017)?

The Keynesian narrative attributes the bulk of business cycles to shifts in aggregate demand. Low aggregate demand is also considered to be an important force behind the Great Recession and the slow recovery from it. In this article, I start by reviewing why this hypothesis finds little place in the Neoclassical, or

The editor in charge of this paper was Dirk Krueger.

Acknowledgments: This article was prepared for the Schumpeter Lecture given at the 2016 Annual Meeting of the European Economic Association. I thank Dirk Krueger for detailed feedback; Daron Acemoglu, Olivier Blanchard, Harris Dellas and Fabrice Collard for comments; 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 is based.

E-mail: angelet@mit.edu

1. See, inter alia, Eggertsson and Krugman (2012); Hall (2011); Mian et al. (2013); Mian and Sufi (2014).

Journal of the European Economic Association
Preprint prepared on 3 March 2018 using jeea.cls v1.0.
RBC, framework and how exactly it is accommodated in the New Keynesian framework. This allows me to highlight, not only the subtlety of the Keynesian narrative, but also some of the challenges that the macroeconomist faces when trying to make sense of the observed business cycles. I next review how my own research has sought to shed new light on the Keynesian narrative, on the mechanisms that drive business cycles, and on related policy questions.

But let me first clarify the philosophy behind this article. As Robert E. Lucas, Jr. once said, 2

“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 usefulness and the empirical validity of the Keynesian narrative, 3 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 idea that low demand can trigger a drop in employment and outcome seems obvious from a partial-equilibrium (PE) perspective, it can be vacuous when considering the general-equilibrium (GE) response to economy-wide shocks. 4 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 are well understood within the field of macroeconomics but less so outside it. They help distinguish the modern macroeconomic paradigm, which builds on microeconomic foundations and spells out how behavior is shaped by incentives, constraints, and expectations, from the old IS-LM framework, which made more ad hoc assumptions on behavior, downplayed the role of expectations, and was severely vulnerable to the Lucas critique (Lucas 1976). They help define the “natural” rate of interest and the “output gap,” central elements of the modern theory of monetary policy. And they help explain the theoretical constructs that macroeconomists often use as proxies for shifts in consumer spending and aggregate demand, such as shocks to the discount factor of the consumers.


3. In an early contribution, Blanchard and Quah (1989) used Structural VARs to provide evidence in support for the idea that business cycles are driven by shifts in aggregate demand. The same idea is corroborated by recent work that exploits the regional variation in business cycles, such as Mian and Sufi (2014) and Beraja et al. (2016), as well as by work that exploits survey evidence, such as Bachmann and Zorn (2018). Note that this evidence is separate from the one regarding either price rigidity or the real effects of monetary policy. The latter kind of evidence is not in dispute, but are also not necessarily related to the points I wish to make here.

4. Henceforth, PE and GE as acronyms for, respectively, partial equilibrium and general equilibrium.
Such shocks should not be taken literally. They help mimic within simple models the effects of, say, a credit crunch on the consumers (Eggertsson and Krugman 2012; Guerrieri and Lorenzoni 2017), an increase in precautionary savings (Challe et al. 2017), or pessimistic beliefs about future income (Lorenzoni 2009). According to the Keynesian narrative, any of these forces translates into a drop in aggregate demand, which in turn triggers a recession, that is, a joint reduction in employment, output, consumption and investment.

Building on these points, Subsection 2.2 reviews why this narrative 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 for consuming 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 most familiar picture from microeconomics, namely that of a downward-sloping demand curve intersecting with an upward-sloping supply curve. It follows that the inconsistency of the Keynesian narrative with the baseline RBC model is proof of how subtle this narrative is. It may appear self-evident from a PE perspective, but it falls apart in a basic, micro-founded GE context.

To salvage the Keynesian narrative, and to make sense of the related evidence, additional assumptions are needed. In the New Keynesian model, the key assumptions are the following. First, there is a constraint in how

---

5. 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 common knowledge of the current state of the economy and common belief of its future prospects.

6. Early attempts to reconcile the Keynesian narrative with micro-founded, rational-expectations settings included the literature on multiple equilibria and sunspot fluctuations, which I discuss in the sequel, and the works of Lucas (1972, 1973) and Barro (1976, 1978) on nominal confusion and unexpected monetary shocks. These works, which built on the Friedman’s notion of the “natural rate” (Friedman 1968), sought to explain why an apparent trade-off between unemployment and inflation—a Phillips 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 exploitable structural relation between inflation and real economic activity, in the form of the New Keynesian Philips curve. I discuss the central position that this relation plays in the theory and some discomforting evidence about it in Subsection 2.5.
fast nominal prices (or nominal wages) can 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-binding; in macro jargon, “monetary policy does not replicate flexible prices”. I fill in the equations that explain the words in Subsection 2.4.

These assumptions are realistic. But they are not part of the demand-and-supply picture that is familiar from microeconomics, underscoring the subtlety of the Keynesian narrative. They also lead to three empirical challenges. First, the model is able to generate realistic business cycles out of shocks to consumer or firm spending only with the help of additional, more dubious, “bells and whistles” (more on this later). Second, the microeconomic data on prices do not necessarily justify the macro-level rigidity that the quantitative versions of the model require in order make sense of the observed business cycles. Third, the model depends on a Philips curve, but the evidence about old and new Philips curves is discomforting, to say the least.

In addition to these empirical challenges, the New Keynesian framework suffers, at least in my view, from an inherent conceptual defect: the only way one is allowed to make sense of demand-driven fluctuations within that framework is, in effect, by equating a drop in aggregate demand with a monetary contraction. By this I mean a contraction relative to the benchmark of replicating flexible-price allocations.

This is at odds with a long tradition that sought to understand recessions as the product of coordination failures and of disruptive GE feedback loops, which can be active even when prices are flexible. See the seminal contributions of Diamond (1982) and Cass and Shell (1983) and the large subsequent literature that captured these ideas in models featuring multiple equilibria and sunspot fluctuations.7 In my view, this literature was after something real and important, which the New Keynesian model assumes away mostly because of the inconvenience of working with multiple-equilibrium models.

This background explains both the empirical motivation behind my work on business cycles and its position in the literature. When I look at the available evidence, I feel compelled to accommodate the concept of demand-driven business cycles while abstracting from nominal rigidity and/or constraints on monetary policy.8 And when I think of the workings of large, decentralized, market-based economies, I gravitate towards theories that give a central position to coordination.

---

8. As already mentioned, the older literature on coordination failures and sunspot fluctuations also tried to rationalize demand-driven fluctuations without nominal rigidities. Recent works aimed at the same goal include Beaudry and Portier (2013), Benhabib et al. (2015) and Bai et al. (2017). Related is also the theory of demand shocks developed in Lorenzoni (2009), although that one hinges on nominal rigidity.
In a series of co-authored papers, I have thus expounded a new theoretical approach, one that shifts the focus from nominal rigidity to frictional coordination. By this, however, I do not mean a coordination failure in the sense of multiple equilibria. Instead, I mean the accommodation of the idea that the agents may have an imperfect, and idiosyncratic, understanding of one another’s behavior, of the current state of the economy, and of its future prospects.

From a game-theoretic perspective, this imperfection amounts to removing common knowledge of the state of the economy and introducing higher-order uncertainty, that is, uncertainty about the beliefs and actions of others. Such uncertainty can be the by-product of either the decentralization of market interactions, along the lines of Lucas (1972), Townsend (1983), and Prescott and Rios-Rull (1992), or of rational inattention and costly contemplation, along the lines of Sims (2003, 2010) and Tirole (2015). One way or another, the key for my purposes is to accommodate the possibility that economic agents are unable to coordinate their expectations and their behavior with the same level of finesse as that presumed in workhorse macroeconomic models.

The considered friction has the flavor of a coordination failure. But whereas the traditional formalization of this notion provided in Diamond (1982), Cass and Shell (1983) and the related literature relies on selecting one among many equilibria, the formalization adopted in my work does not. It can therefore be embedded in models featuring a unique equilibrium, including the textbook versions of the RBC and New Keynesian frameworks. And it can be thought of as an endemic feature of any realistic, decentralized, market-based economy.

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, I introduce a new kind of shocks, which cause extrinsic shifts in higher-order beliefs and ultimately rationalize waves of pessimism and optimism about the short-term economic outlook (Angeletos and La'O 2013; Angeletos et al. 2015). In the other, I refrain from the introduction of such shocks and, instead, show how the considered friction modifies the propagation of familiar forms of demand shocks (Angeletos and Lian 2017b).

9. Note that common knowledge of an event or a property \( X \) means, not only that everybody knows \( X \) (first-order knowledge), but also that everybody knows that everybody knows \( X \) (second-order knowledge), that everybody knows that everybody knows that everybody knows \( X \) (third-order knowledge), and so on.

10. Such “micro-foundations” of higher-order uncertainty are interchangeable for my purposes. I adopt the first approach in Angeletos and La’O (2013) and Angeletos and Lian (2017a), the second in Angeletos and La’O (2011) and Angeletos and Sastry (2017), and a mixture in Angeletos and Lian (2016a). See also Angeletos et al. (2015) and Angeletos and Lian (2017a) for how similar goals can be achieved by relaxing the rational-expectations and common-prior assumptions.
Either way, the Keynesian narrative is disconnected from nominal rigidity. 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 New Keynesian model by documenting that expectations of inflation, as measured in surveys, did not turn negative during the Great Recession. But this leaves unanswered the question of why expectations did not move as the model predicts. For an attempt to offer a complete explanation within the New Keynesian framework, see Christiano et al. (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 because the formalizations of demand-driven fluctuations described above are consistent with—albeit not dependent on—nominal rigidity, but also because of the following two reasons. First, the considered friction can modify the propagation mechanisms and the policy predictions of the New Keynesian framework in empirically appealing ways. Second, the same friction can offer a micro-foundation for some of the more dubious bells and whistles of that framework. These points define 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 et al. 2012; McKay et al. 2016). But a variant of this puzzle appears to apply more generally: the macroeconomic times series indicate small and sluggish responses of both inflation and the real quantities to a variety of identified shocks.\footnote{For instance, Gali (1999) documents a sluggish response of real quantities to identified technology shocks, while Christiano et al. (2005) document a sluggish response of both the real quantities and inflation to identified monetary shocks. See also the complementary discussion in Sims (2003).} What is more, the sluggishness in the response of the actual outcomes appears to be accompanied by an even larger sluggishness in the response of expectations (e.g., Coibion and Gorodnichenko 2012, 2015a) and the adjustment friction appears to be larger at the macro level than at the micro level (e.g., Havranek et al. 2017).
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 generate realistic business cycles and to match the macroeconomic data, such as certain forms of habit in consumption and adjustment costs in investment, or the so-called hybrid version of the New Keynesian Philips curve (Angeletos and Huo 2018).

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, 2017a). 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.

**Remark 1.** Informational frictions have interesting effects even in a single-agent decision context. 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, which followed the contributions of Morris and Shin (1998, 2001, 2002, 2003) and Woodford (2003). For a synthesis of these literatures and additional references, I refer the reader to my chapter in the Handbook of Macroeconomics (Angeletos and Lian 2016b). Complementary is also the literature on rational inattention (Sims 2003, 2010).

**Remark 2.** Informational frictions can rationalize monetary non-neutrality even in the absence of sticky prices and menu costs. This an important point, which goes back to Lucas (1972, 1973) and Barro (1976, 1978) and has been revisited by Mankiw and Reis (2002), Woodford (2003), 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.

![Figure 1. Demand Shocks in Partial Equilibrium](image)

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 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 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, a standard practice 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; that is, the relevant $P$ from a GE perspective is the real interest rate. By the same token, a shock to aggregate demand means a shock in the demand for goods today relative to the demand for goods.\footnote{This GE perspective differs from that of the IS-LM framework, which represents aggregate demand (AD) and aggregate supply (AS) as functions of the nominal price level. Friedman (1968), Lucas (1972, 1973) and Barro (1976, 1978) sought to make sense of a AD-AS picture in terms of an expectations-augmented Philips curve, in effect by interpreting the variable in the vertical axis as the gap between the actual and the perceived, or expected, nominal price level. The New Keynesian framework blends the IS-LM and the GE perspectives by emphasizing the role of forward-looking behavior and the importance of the real interest rate on the demand side, while representing aggregate supply in terms of the New Keynesian Philips Curve. In this article, I prefer to stick with the “pure” GE perspective that looks at aggregate demand and aggregate supply in terms of the real interest rate, rather than the price level or inflation, even when I allow for nominal rigidity.}

How should we think about a “drop in aggregate demand” from this GE perspective? To be concrete, consider the drop in consumer spending during the Great Recession. The evidence in Mian et al. (2013) and Mian and Sufi (2014) suggests that this drop was 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 changes to consumer-credit conditions is quite common in the macroeconomics literature. See, for example, Eggertsson and Woodford (2003), Werning (2012) and Christiano et al. (2015). For more realistic treatments, see Eggertsson and Krugman (2012), Guerrieri and Lorenzoni (2017), and Challe et al. (2017).}

Let us take such a shock for granted and let us ask how it 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. A minimalistic GE model of this kind can be found in Barro (1997), Angeletos and Lian (2017b), and elsewhere.

The employed model is deliberately simple. Having two periods is of essence so that we can talk meaningfully about shifts in aggregate demand, that is, in the demand for goods today relative to tomorrow. Having both a consumption-leisure and a consumption-saving choice in the first period allows the model to shed light on whether demand and supply shocks can trigger comovement in employment, consumption, and investment. Abstracting from such choices in the second period is only for simplicity.

In the sequel, I study two versions of this simple model. In the first, 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 rigid 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

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

(1)

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 \( U(c, \ell) = u(c) + v(\ell) \), where \( u \) and \( v \) are strictly increasing and strictly concave. I also impose \( \ell_2 = \bar{\ell}_2 \), for some exogenous \( \bar{\ell}_2 \), so that 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,

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

(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.15 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 \( i_1 = k_2 - k_1 \) 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, a lower \( \theta \) represents a negative demand shock. By contrast, a lower \( A \) can be interpreted as a negative supply shock. The question addressed in the rest of this section is how the first-period equilibrium outcomes respond to each of these shocks.16

As long as the First Welfare Theorem applies (which is the case here, 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, and the prices that support the planner’s solution as part of a competitive equilibrium, are given by the following:

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

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

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

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).

17. The second condition does not contain an expectation operator because I have abstracted from uncertainty.
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 $A$. This triggers a temporary drop in output and income. Because of the desire to smooth consumption, this is absorbed partly by a drop in consumption and partly by a drop in saving and hence in investment. Provided the substitution effect on labor supply dominates the wealth effect, which is the empirically relevant case and the case that I assume here, employment also falls. This effectively reviews the RBC model, in which recessions are explained by adverse supply shocks.

What about a negative demand shock, that is, a drop in $\theta$? When $\theta$ falls, consumers have an “urge for saving,” 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.\(^{19}\)

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

\(^{18}\) 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.

\(^{19}\) 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 et al. 2007).
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.

The equilibrium adjustment is illustrated in Figure 2. This figure is the GE analogue of Figure 1. On the horizontal axis, we have the aggregate quantity of goods today. On the vertical, we have their relative price, namely the real interest rate. The two AS curves represent the aggregate supply of goods before and after the shock; the two AD curves represent the corresponding aggregate demands.

![Figure 2. Demand Shocks in General Equilibrium](image)

For given $\theta$, the AS curve is obtained as follows. First, combine the optimal labor-supply condition of the household and the Euler condition to get labor supply as a function of the real wage and the real interest rate. Next, note that labor supply increases with the real interest rate because of intertemporal substitution. Finally, use the optimal labor-demand condition of the firm to solve out for the real wage. This gives the quantity of labor that clears the labor market and the resulting supply of goods as increasing functions of the real interest rate.

The AD curve, on the other hand, is given by the sum of the firm’s demand for investment and the household’s demand for consumption. The former is decreasing in $r$ due to the diminishing marginal product of capital. The latter obtains from the household’s Euler condition and is decreasing in $r$ because of intertemporal substitution.

What happens as $\theta$ falls from $\theta_1$ to $\theta_2$? Holding $r$ constant, the demand for investment remains unaffected, but the demand for consumption falls. This leads the AD curve to shift left. At the same time, because the urge for saving
(or the tighter consumer credit) stimulates labor supply, the AS curve shifts right. All in all, the equilibrium \( r \) falls, from \( r_1^* \) to \( r_2^* \), and the equilibrium \( y \) increases, from \( y_1^* \) to \( y_2^* \).

This further clarifies why the PE logic is misleading: in GE, the negative demand shock is, in effect, also a positive supply shock, due to the aforementioned labor-supply response. One can devise a variant in which the stimulating effect of the \( \theta \) shock on labor supply is shut down while its contractionary effect on the demand for consumption is preserved.\(^{20}\) Even in that variant, however, investment would move in the opposite direction than consumption, implying that the model cannot generate a realistic recession in response to the shock.\(^{21}\)

To sum up, making sense of the Keynesian narrative requires one to move beyond the most basic principles of microeconomics: one must 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 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, TFP, 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 I 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 et al. (2017): that paper develops an extension of the RBC model that adds search frictions in commodity markets and lets preference shocks generate endogenous movements in measured TFP.

To understand what I 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

---

20. For instance, suppose there are two groups of households, called 1 and 2, and suppose that the only group 1 supplies labor in the first period. Suppose next that that the \( \theta \) shock affects only group 2. Then, the shock affects aggregate demand by affecting the consumption of group 2, but does not affect aggregate supply.

21. Unless, of course, the model is augmented with some other mechanism—such as pessimistic beliefs about the future or aggravated financial frictions on the side of the firms—that causes the demand for investment to fall in tandem with the demand consumption.
planner’s intra-temporal condition with the following variant:

$$\frac{u'(\ell_1)}{u'(c_1)} = (1 - \tau_1) AF_n(k_1, n_1),$$

(7)

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.\textsuperscript{22} 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.\textsuperscript{23} 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.

This is also a key challenge of a recent line of work that generates business cycles out of “uncertainty shocks” (shocks to second moments) in extensions of the RBC model that allow for firm heterogeneity and adjustment costs (Bloom 2009; Bloom et al. 2012). These models hinge on generating strong co-movement between aggregate employment and aggregate TFP, a property that is at odds with the data.

\subsection*{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, nominal rigidity. Firms are price-setters, rather than price takers, and adjust their nominal prices only periodically. 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.

\textsuperscript{22} 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 $i_1 (= k_2)$.

\textsuperscript{23} 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.
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 et al. 2008). But even when such a policy is sub-optimal, or just infeasible, the benchmark defined by it is instrumental for understanding whether and how the New Keynesian model can accommodate the desired narrative.

Before proceeding, let me reiterate that my discussion of the New Keynesian model in this subsection, just as my review of the RBC model in the previous subsection, is centered around the questions of "what are the equations that explain the words" or "how the model works." This may sound less exciting than the question of "how the real world works." But the latter question is ill defined, for it is only through the endless back-and-forth between models and data that we develop a meaningful understanding of how the real world works. This explains my obsession with the precise manner in which the New Keynesian model accommodates demand-driven business cycles.

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.

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:

\[
\frac{v'(\ell_1)}{w'(c_1)} = w = \frac{1}{\mu_1} AF_n(k_1, n_1) \tag{8}
\]

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

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

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. In particular, \(\mu_1\) plays the same role as the labor wedge in Subsection 2.3, while \(\mu_2\) shows up as a wedge in the Euler condition. 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.\textsuperscript{24}

When prices are flexible, $\mu_t$ can be 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 fixed $\mu^* > 1$. 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 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. 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 $v'(1 - c_1/A)$. Finally, suppose that the monetary authority guarantees that the inflation rate in the second period is zero, which

\textsuperscript{24} I have suppressed two equilibrium conditions, which are not central for the arguments made in this subsection but explain how monetary policy controls the realized markups. 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 the New Keynesian Philips Curve, a structural equation that relates inflation to real output. 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 its policy instrument, the nominal interest rate.
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 \( \mu_1 \).²⁵,²⁶

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.

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. As illustrated in Figure 2, 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 \( \mu_1 \) increase as \( \theta \) falls. This happens when the monetary authority does not allow the actual interest rate, \( r \), to fall as much as its flexible-price counterpart, \( r^* \). For instance, consider Figure 2 and suppose that, as \( \theta \) falls from \( \theta_1 \) to \( \theta_2 \), the monetary authority keeps the real interest rate constant at the pre-shock natural level, \( r^*_1 \). Then clearly investment does not change and output, which is now demand-determined, falls by exactly the same amount as the demand for consumption.

In preventing the real interest rate from adjusting to \( r^*_2 \), 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 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

---

²⁵. I 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. This is easier to address in the full-blown version of the New Keynesian model.

²⁶. Clearly, 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 \( \mu^* \).
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.\(^{27}\)

### 2.5. Summary and 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, formalized as a surge in the desire to save, 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 the aforementioned 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 New Keynesian model and an additional prediction that follows from these assumptions and is hard to reconcile with the data.

Let me first focus on the assumptions. The assumption that prices are sticky or, more broadly, that there is some kind of nominal rigidity is supported by the available micro-economic evidence (Klenow and Malin 2010; Nakamura and Steinsson 2008). 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 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.\(^{28}\)

---

27. 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.

28. For the aggregate implications of menu costs, see also Caballero and Engel (1993, 2007).
More recently, Golosov and Lucas (2007) corroborated the empirical relevance of the above lesson by quantifying the monetary non-neutrality implied by a basic menu-cost model calibrated to the available microeconomic evidence and by showing this was significantly smaller than the one typically assumed in the New Keynesian framework. Subsequent work by Gertler and Leahy (2008), Nakamura and Steinsson (2010), Midrigan (2011), Alvarez and Lippi (2014) and others has sought to qualify or overturn Golosov and Lucas’s conclusions, highlighting how the implied level of monetary non-neutrality depends on “details” such as the precise stochastic properties of the idiosyncratic shocks to production costs or the number of products sold by each firm. All in all, however, it seems fair to say that the debate on the quantitative importance of nominal price rigidity has not been settled. The 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 for validating the New Keynesian formalization of demand-driven fluctuations. This formalization 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, these gaps are not 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:

$$\pi_t = \kappa x_t + \beta E_t[\pi_{t+1}],\quad (10)$$

or, by iterating,

$$\pi_t = \kappa E_t \left[ \sum_{k=0}^{\infty} \beta^k x_{t+k} \right],\quad (11)$$

where $\pi_t$ denotes the inflation rate between period $t-1$ and period $t$, $E_t[\cdot]$ 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 $\mu_2 = \mu_2^* \text{ and } x_2 = 0$, 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 ought 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.29

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 NKCP. That evidence 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.30 But even if one takes for granted this assumption, the empirical relation between the labor share and inflation is weak.31

---

29. 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.

30. A closely related approach is taken in Sbordone (2002).

31. For instance, Angeletos et al. (2017) use a Structural VAR approach to document that, 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. See also the complementary critique in King and Watson (2012).
One can try to salvage the NKPC by arguing that most of the variation in inflation is driven, not by variation in aggregate demand, but 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 \( \kappa \mu_t^* \). 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 \( \pi_t \). In line with this observation, the extant DSGE literature (e.g., Smets and Wouters 2007) attributes almost the entirety of the inflation fluctuations to implausible markup shocks.

It is hard to view the above as empirical validation of the New Keynesian formalization of demand-driven fluctuations. One may counter-argue that I am taking the NKPC and its descendants 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. I am quite sympathetic to this view. Part of the work that I describe in Section 4 indeed seeks to fix some of the empirical failures of the New Keynesian model and especially of the NKPC. However, because the New Keynesian model equates demand-driven business cycles to gaps, which in turn are tied 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.\(^{32}\)

To sum up, when I digest the lessons of the literature, look at the aggregate time series, or try to understand why the Great Recession was not the Great Deflation, I feel compelled to move beyond the New Keynesian framework. And while I recognize the evidence about monetary non-neutrality, as documented in the literature or apparent in episodes such as the Volker disinflation, I am not convinced that nominal rigidity is the most essential reason for why demand shocks can trigger business cycles. Instead, like Beaudry and Portier (2013) and the older tradition that emphasized the role of coordination failures, I believe that the data demand a theory of non-inflationary demand-driven fluctuations. I review how my research offers such a theory in the next section.\(^{33}\)

---

32. See also the review of the empirical literature on the NKPC in Mavroeidis et al. (2014) and the topical discussion in Blanchard et al. (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.

33. Throughout this subsection, I have quibbled with the New Keynesian framework, and especially the NKPC, in order to motivate the alternative theory reviewed in the next section. However, the friction I am concerned with also offers two modifications of the
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.

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. 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. 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

---

34. See, inter alia, Diamond (1982); Cass and Shell (1983); Cooper and John (1988); Guesnerie and Woodford (1993); Woodford (1991); Benhabib and Farmer (1994, 1999).
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 prevents such a perfect consensus by introducing heterogeneous information and allowing the agents to face non-trivial higher-order uncertainty. This approach, which builds heavily on Morris and Shin (1998, 2002, 2003) and Woodford (2003), helps accommodate the notion of coordination failure in models that admit a unique equilibrium. 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). 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 the door to forces that resemble 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.

35. Nevertheless, 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. This follows from essentially the same argument as the one reviewed in the next subsection. See also the complementary discussions in Morris and Shin (2002) and Angeletos and Werning (2006) about the sunspot-like function of public signals in environments in which coordination is important.
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 microfoundations. 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 latter’s output, 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 together 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

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

(13)

In short, \( i \)’s output depends, not only on her own productivity and on her belief of the beliefs of her partner.
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

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

and similarly for \(y_{jt}\). And since this is true for every matched pair \((i,j)\), aggregate output is given by

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

where \(A_t\) denotes aggregate TFP. In short, the complete-information version of the model under consideration is a close cousin for the RBC model: it attributes the business cycle to supply shocks (in the form of aggregate TFP 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 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

\[
x_{it}^1 = A_{jt} + \varepsilon_{it},
\]

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

\[
x_{it}^2 = \varepsilon_{jt} + \xi_{it},
\]

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 a 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, $\xi_t$ captures the sentiment (belief) that a farmer has about her terms of trade and the returns to labor. 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 technology and given marginal costs, these fluctuations can be said to have a Keynesian flavor: they feel and look like demand-driven fluctuations. Keep in mind, though, that the separation between demand and supply is always fussy in GE settings with flexible prices. As a negative sentiment shock causes each farmer to produce less of his own good and to demand less of the goods of others, supply and demand move together and indeed feed one another. In terms of Figure 2, it is therefore as if there is a joint, and self-reinforcing, leftward shift in both the AS and the AD curve.

### 3.2. Sentiments and the Business Cycle

A few recent paper expand on the ideas described above, or push forward closely related theories of the business cycle. Angeletos et al. (2015) and Huo and Takayama (2015) develop more realistic versions of Angeletos and La’O (2013) and confront them with the data. Wu and Young (2017) study an extension that introduces an asset market and argue that this kind of model can help explain jointly the business cycle and the volatility in asset markets. Sockin and Xiong (2015) consider an application to commodity markets and use it to explain certain empirical puzzles. Benhabib et al. (2015) and Acharya et al. (2017) show that the signal-extraction problem that the firms face between idiosyncratic and aggregate shocks can open the door to a similar type of sentiment-driven fluctuations as that described above, even in the absence of exogenous shocks to higher-order beliefs. Chahrour and Gaballo (2017) offer a complementary formalization that allows the belief waves to obtain from small shocks to technology or other fundamentals and ties them to expectations of wealth.
Ilut and Saijo (2016) and Pei (2018) engineer similar kinds of belief waves from the combination of idiosyncratic uncertainty, ambiguity, and learning.

Evaluating the quantitative potential of this class of models can run quickly to computational challenges. A joint project with Fabrice Collard and Harris Dellas bypasses these challenges and develops a 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 et al. 2015), which I review next, 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. The method, however, is more general and can be used for other purposes, too. For instance, it can readily capture the sluggish response of higher-order beliefs to shocks in monetary policy, TFP, and other fundamentals, as in the works of Woodford (2003), Nimark (2008, 2017), Mackowiak and Wiederholt (2009, 2015), and Angeletos and La’O (2010).36

Consider 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.37

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.

---

36. The developed method can also capture the kind of myopia, and extra discounting of the future, that Angeletos and Lian (2016a) rationalize with a relaxation of the common-knowledge properties of the New Keynesian model and which Gabaix (2016) and Farhi and Werning (2017) replicate with appropriate departures from rational expectations.

37. The baseline model considered in Angeletos et al. (2015) differs from the one considered in Angeletos and La’O (2013) and reviewed before in three respects. First, there is capital accumulation so that the model can speak to the co-movement of employment, consumption, and investment. Second, instead of pairwise matches among farmers, there is the standard interaction of multiple households and firms in labor and capital markets. Third, the relevant belief waves are engineered with the help of heterogeneous priors. The essence, however, is the same.
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 changing the long-term prospects. 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 following key regularities of the US macroeconomic times series: 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 et al. (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 $t$:

\[
\frac{v'(\ell_t)}{u'(c_t)} = (1 - \tau^\ell_t) A_t F_n(k_t, n_t),
\]

\[
u'(c_t) = \beta \mathbb{E}_t \left\{ u'(c_{t+1}) \left[ 1 + (1 - \tau^k_{t+1}) A_{t+1} F_k(k_{t+1}, n_{t+1}) \right] \right\},
\]

where the $\tau$’s capture the endogenous wedges induced by the sentiment shock. In particular, a negative sentiment shock manifests as a joint increase in $\tau^\ell_t$ and $\tau^k_t$: pessimistic beliefs about the choices of others are akin to a joint tax on labor and capital.$^{38}$

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 et al. (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.

Complementing these findings, Angeletos et al. (2017) and Levchenko and Pandalai-Nayar (2017) provide empirical evidence, based on Structural VARs, that the bulk of the observed business cycles is consistent with a type of fluctuations like the one captured in my work and in the literature cited in the beginning of this subsection—and unlike the one captured in competing models that emphasize technology, monetary, news, or uncertainty shocks.$^{39}$

$^{38}$ This follows from Subsection 5.4 of Angeletos et al. (2015), which discusses how sentiment shocks manifest as wedges in terms of business-cycle accounting (Chari et al. 2007). Related theories of belief-induced wedges appear in Ilut and Saijo (2016) and Bhandari et al. (2016).

$^{39}$ Even if one does not embrace the formalization of sentiment- or confidence-driven fluctuations that my work and the related literature has put forward, there seems to be
Finally, it is worth emphasizing that the variation in confidence does not have to be extrinsic. Angeletos and Lian (2017b) and Ilut and Saijo (2016) consider models that feature similar kinds of belief waves, and of belief-driven wedges, as the ones discussed above, except that the beliefs and the associated wedges are allowed to covary with more conventional structural shocks, such as financial or discount-factor shocks. This helps explain why a drop in confidence may be triggered by an adverse financial shock, while a boost in confidence may be accomplished by a fiscal stimulus.

### 3.3. From Shocks to Propagation

So far I have tried to make sense of demand-driven business cycles using an extrinsic sentiment shock. As already noted, 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 2017b). We argue that removing common knowledge of the kind of consumer-spending or discount-rate shocks described earlier, and allowing such shocks to be confounded with idiosyncratic shocks in firm profitability and household income, permits these shocks to generate realistic business cycles within the RBC framework. We further show how the same ingredients give rise to a feedback mechanism that resembles the Keynesian multiplier and that helps rationalize large fiscal multipliers in the short run, despite the absence of any kind of nominal rigidity. We finally discuss how this provides a possible micro-foundation and an appealing re-interpretation of the kind of belief waves described earlier.

Distinct but complementary theories that allow the level of “confidence” to vary endogenously in response to conventional forms of demand or supply shocks are developed in Chahrour and Gaballo (2017), Ilut and Saijo (2016), and Schaal and Taschereau-Dumouchel (2015). The first focuses on the feedback between prices and wealth; the second on the interaction of ambiguity and learning; the third on non-convexities in production. In the rest of this section, I focus on the mechanism formalized in Angeletos and Lian (2017b), not only a broader take-home lesson. The quantitative explorations of Angeletos et al. (2015) and Huo and Takayama (2015), the VAR-based empirical evidence in Angeletos et al. (2017) and Levchenko and Pandalai-Nayar (2017), and the complementary evidence in Beaudry and Portier (2013) and Beaudry et al. (2015) all point in the same direction. There are important regularities in the aggregate time series for which the more conventional business-cycle literature has failed to provide a convincing parsimonious explanation.
because this is part of my own work, but also because of its tight connection to the Keynesian narrative, which is the unifying theme of this lecture.

The baseline model in Angeletos and Lian (2017b) 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 realization of it causes a boom or a recession depends critically on the whether the shock is common knowledge or not.

When information is complete and the shock is common knowledge, 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 once there is incomplete information and lack of common knowledge. The discount-rate shock triggers a drop in the demand faced by each farmer (or firm). Because information is incomplete, some farmers may not be able to tell whether the drop in their demand is due to idiosyncratic or aggregate reasons. As a result, these farmers may work less and may instruct their sibling-consumers to spend less. But as these latter 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.

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 be 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, or slows down, GE effects by arresting the adjustment of the expectations of the actions of others to aggregate shocks. I first articulate the basic idea within an abstract framework. 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 the same idea offers an empirically plausible micro-foundation of some of the “bells and whistles” of the New Keynesian model.

4.1. Dampening General Equilibrium

The basic idea that lack of common knowledge attenuates GE effects was articulated in Angeletos and Lian (2017a) within the context of an elementary, and decentralized, Walrasian economy. There are two periods, “today” and “tomorrow.” There are three goods. 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 the 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.

These assumptions are deliberately stark. They nevertheless help 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.

How does the considered 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.

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 have 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. This interval corresponds to the interval between points Y and Z in Figure 4. By imposing rational expectations together with common knowledge of the shock, standard modeling practices pick, perhaps inadvertently, the upper bound of this interval, namely point Z. 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 the GE adjustment of the modified economy can lie anywhere along the interval between Y and Z.

The logic 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 underlying shock, 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 the practice of combining rational expectations with common knowledge of the underlying shocks “overstates” the importance of GE mechanisms. Regardless of whether one considers a setting in which the GE effects of a shock amplify its PE effects (left panel in the figure) or a setting in which the opposite is true (right panel), removing common knowledge of the shock is akin to dulling the GE mechanism.

4.2. Application: Forward Guidance and Fiscal Stimuli

I now discuss how the aforementioned ideas shed new light on the question of how policy influences aggregate economic activity. In particular, consider the New Keynesian framework and ask how monetary policy influences aggregate demand within that framework. By changing the interest rate(s) faced by the typical borrower or saver, monetary policy has a direct effect on the budget, the incentives, and the behavior of that agent, even when the behavior of all other agents remains unchanged. This kind of PE (or, more precisely, decision-theoretic) effect is relatively modest. The largest part has to do with GE mechanisms, which stem from the response of all other agents in the economy and 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 ideas described in the previous subsection, my work on “Forward Guidance without Common Knowledge” (Angeletos and Lian 2016a) shows how removing common knowledge of the monetary policy and of its consequences is akin to attenuating all the aforementioned GE effects. As a result, the response of the economy to forward guidance and to any other news about the future is reduced.

What is more, this attenuation increases with the horizon that the agents have to forecast. This is because policies, or shocks, that work over long horizons map to beliefs of high order and are therefore more sensitive to any given imperfection in information and coordination. It is therefore as if the economy is populated by a representative agent who is myopic and discounts the future more heavily than what it is rational.40

These insights help resolve the so-called forward guidance puzzle (Del Negro et al. 2012; McKay et al. 2016). The latter refers to the New Keynesian model’s 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 driven by the GE effects described above and, in particular, by the property that these GE effects pile up as the horizon increases. By dulling these effects, my work brings the predictions of the model closer both to the available evidence and to the appealing PE logic that interest rates at long horizons ought to have small effects due discounting.

These insights 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, my work indicates that fiscal stimuli should be front-loaded in order to minimize the bite of frictional coordination. Angeletos and Lian (2016a) illustrate this point within the ZLB context. In ongoing work, we extend the analysis away from the ZLB and examine how the proposed mechanism interacts with the one based on “hand-to-mouth consumers” (Galí et al. 2007, Kaplan and Violante 2014) and short horizons (Del Negro et al. 2012).

40. Gabaix (2016) and Farhi and Werning (2017) accommodate similar forms of myopia by dropping rational expectations.


4.3. Myopia and Anchoring, with application to the NKPC

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 2018) shows how incomplete information can offer a micro-foundation for some of the more dubious “bells and whistles” that the New Keynesian framework requires in order to generate realistic business cycles, such as the so-called hybrid version of the NKPC or specific kinds of adjustment costs in investment and habit formation in consumption.

To illustrate, consider the NKPC. Its standard version is given by

\[ \pi_t = \kappa x_t + \beta E_t [\pi_{t+1}] , \]

where \( \pi_t \) denotes inflation, \( x_t \) denotes the output gap (or 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 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:

\[ \pi_t = \kappa x_t + \beta \delta E_t [\pi_{t+1}] + \gamma \pi_{t-1} , \]

where the scalars \( \delta, \gamma \) are pinned down by the underlying parameters \( \kappa, \beta \) and the information structure. These scalars satisfy \( \delta < 1 \) and \( \gamma > 0 \). It is therefore as if the firms discount the future more heavily and, in addition, anchor their optimal reset prices to the past price level.
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 learning: 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.\(^{41}\)

These features—discounting of the future and anchoring to the past—induce the kind of empirical patterns that the DSGE literature have sought to capture with a variety of bells and whistles, such as the hybrid NKPC and the specific forms of habit in consumption and adjustment costs in investment popularized by Christiano et al. (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 suggests that the same objectives can be accomplished with one friction. What is more, while the existing approach requires the relevant adjustment friction—say, habit formation—to be equally present at the micro and the macro level, my approach explains why the friction appears to be much larger when looking at the macroeconomic data than when looking at the microeconomic data (Havranek et al. 2017). Last but not least, the proposed approach is 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).

Let me close this section with the following remark. Throughout, I have focused on the implications of adding informational frictions in the New Keynesian framework while maintaining the usual formalization of nominal rigidity. The works of Woodford (2003), Mankiw and Reis (2002), Mackowiak and Wiederholt (2009) and Chung et al. (2015) make complementary but different contributions by showing how informational frictions may substitute for Calvo-like sticky prices as sources of nominal rigidity. Finally, for additional work on the usefulness of introducing higher-order uncertainty in the New Keynesian framework, see Nimark (2008) and Wiederholt (2016).

\(^{41}\) Woodford (2003) 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 kinds of departure from rational expectations, 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 illustrate the value of a growing literature, of which my work is a small part. This is the recent literature on incomplete information and higher-order uncertainty that was spurred by the influential contributions of Morris and Shin (1998, 2001, 2002, 2003) and Woodford (2003) and that is surveyed in Angeletos and Lian (2016b). All in all, I hope to have conveyed the following three broader lessons.

First, the familiar narrative that a drop in aggregate demand can cause a recession is much subtler than what the related partial-equilibrium logic suggests. By addressing this narrative in both old and new ways, I hope to have raised the reader’s appreciation of its subtlety and to invite further research into its precise meaning.

Second, higher-order uncertainty is a useful modeling device for reconciling coordination failure with equilibrium uniqueness and for accommodating seemingly irrational or self-reinforcing waves of optimism and pessimism. This accommodation can help explain salient features of the data and shed new light on the sources and the propagation of business cycles.

Finally, augmenting rational-expectations models with higher-order uncertainty helps reveal their “robust” or “true” predictions. By this I mean that the accommodation of higher-order uncertainty allows the analyst to disentangle the rational-expectations hypothesis, with its well-known methodological advantages, from the far less palatable assumption that the agents have common knowledge of the current state of the economy, can reach a perfect consensus about its future prospects, and can effortlessly and instantaneously coordinate their responses to policy shifts or other impulses. This disentangling can help resolve certain empirical puzzles and offer new guidance to policy.

References


Journal of the European Economic Association
Preprint prepared on 3 March 2018 using jeea.cls v1.0.


*Journal of the European Economic Association*  
Preprint prepared on 3 March 2018 using *jeea.cls* v1.0.


