A Unified Theory of Growth, Cycles and Unemployment - Part II: Business Cycles and Unemployment

Pollak, Andreas

University of Saskatchewan

14 December 2022

A Unified Theory of Growth, Cycles and Unemployment Part II: Business Cycles and Unemployment*

Andreas Pollak†

Department of Economics, University of Saskatchewan

December 14, 2022

Abstract

The growth framework presented in part I of this paper is extended to include labour market frictions, resulting in a model that has interesting cyclical properties, including the following:

(1) In response to investment-reducing shocks, the model endogenously creates recessions followed by drawn-out recoveries, closely resembling time series data on unemployment, output, investment and asset prices; (2) recessions are fully explained as periods during which frictions prevent instantaneous reallocation, resulting in (3) stock market crashes at the beginning of recessions; (4) the persistently elevated unemployment following recessions is explained as a result of investment dynamics; (5) the model incorporates a mechanism that strongly amplifies investment-reducing shocks while dampening investment-increasing shocks; this leads to (6) a pronounced asymmetry in business cycles, even for symmetric shocks; (7) the model further explains why output can be above trend during investment booms; (8) cyclical fluctuations can be triggered by a variety of shocks, including for example productivity or financial shocks; (9) the model is capable of expectation-driven cycles: the anticipation of future changes can trigger investment booms or recessions without the need for any contemporary productivity changes; (10) the shape of recessions and recoveries is largely driven by model mechanics, and does not rely on particular characteristics of the shock; (11) the model matches the usual cyclical correlations as well as typical RBC models, and in addition to that replicates the skewness of cyclical variables; (12) the model is simpler than most alternative business cycle frameworks and more robust with regards to its reliance on household characteristics for cyclical patterns.

Keywords: business cycles, unemployment, asset prices.

JEL Classification Numbers: E3, J6, G1

*Please check at `http://usask.pollak.org/?research` if an updated version is available. Note that there are two parts to this paper.

†Address: 9 Campus Dr., Saskatoon, SK, S7N 0P9, Canada; telephone: +1 (306) 966-5221; e-mail: *a.pollak@usask.ca*.

1 Introduction

In part I of this paper, we developed a model of endogenous growth that is consistent with a stable growth rate and constant returns to scale at the aggregate level, while being robust with respect to trends in R&D effectiveness, changes in demographics or market power. In this second part, we will extend our growth model to allow for the study of business cycle patterns.

1.1 Plan of this Paper

The following section 2 introduces a business cycle version of the growth framework presented in part I of this paper. The new ingredient that enables us to obtain interesting responses to exogenous shocks is simple search frictions in the labour market. After developing a basic version of the model and briefly discussing its properties, we will investigate the ability of the model to generate realistic business cycles, most importantly recessions followed by prolonged recovery periods. We will be using a calibrated version of the model for this purpose and simulate the responses to shocks. Most of the simulations will be based on generic reallocation shocks. We will also evaluate a number of alternative shocks and disturbances frequently used in the macroeconomic literature with regards to their ability to generate cyclical fluctuations.

In section 3, we will look at how our model can provide new insights into longstanding issues in macroeconomics and finance, such as excess volatility puzzle and the equity premium puzzle. Finally, section 4 summarizes our findings, discusses possible extensions, makes remarks on how our results generalize and compares our framework to popular models in the business cycle literature.

2 Business Cycles

In what follows, we discuss the cyclical properties of our model. The overall setting is very similar to the one used in part I of this paper. Any references to the first part will be prefixed with a Roman numeral I.[1] With a single relatively minor modification, the model will be capable of generating responses to certain shocks that closely match the cyclical patterns observed in aggregate data. This includes a phenomenon that resembles a recession followed by a prolonged recovery period.

Before proceeding, it is important to stress that projects, the units of production introduced in part I, are specific to their industry and sector. Each project employs capital, including knowledge capital, that enables it to be competitive in producing a particular type of good. If the relative demand for the output of different industries changes, the response may involve shutting down projects at a higher rate in the industry that faces reduced demand, while reducing the termination rate or even starting up new projects where demand is increasing. This notion in itself is entirely innocuous in the frictionless setting we have been using so far. Since we assumed that projects can be created or shut down at any time and without cost, there is no real difference between a project being repurposed to produce a different good versus the project being shut down and its factors used to start up a new one in a different industry in the same period. The assumption only becomes important in combination with the labour market frictions we introduce now.

From now on, we will assume that if labour is unemployed, it takes some time for it to become available again to be reemployed. The reasons for being unemployed include being new to the labour force, job separation for idiosyncratic reasons and the termination of the project the unit of labour was assigned to. This last source of unemployment interacts with the industry specificity of production units. Suppose the demand for output in an industry drops. Projects are being shut down, and the previously used production factors are released. While labour becomes unemployed and unavailable for immediate reemployment,

[1] E.g. equation (I-2) or section I-2.2.

the capital is put on the market, where it competes with newly produced investment goods. This temporary increase in the supply of second-hand capital goods reduces the demand for the output of industries in the investment sector, possibly triggering further adjustments and project terminations. This can create an amplification effect that is capable of turning relatively small disturbances into recessions.

2.1 The Firm Sector

The model of the firm sector is the same as in the growth section I-2.2 in part I of the paper. We only need to specify the functional form of ψ and the nature of the shocks affecting production.

We will discuss the effects of a variety of different macroeconomic shocks in subsection 2.6. For the following exposition, we will consider two types of shocks. The first one is a simple TFP shock $\sigma^{[D]}$ that temporarily affects productivity, so that output is given by $y = (1 + \sigma^{[D]})Dk^{\alpha}n^{\beta}$.

The choice of the second shock is motivated by two considerations. First, as seen in part I, firms continuously adjust their production techniques, increasing the factor inputs as the economy grows. New, more productive factor input combinations are explored, and growth acceleration costs (GACs) are expended to implement them. It seems plausible that in this process, firms may reach a point where it becomes clear that the production function will look a little different going forward. For example, the R&D expenditure required to implement a given productivity improvement could be higher than in the past, which would translate into relatively lower productivity of the capital that embodies such improvements. In the language of figure I-1, as firms move though the space of production techniques towards the top right, they may encounter isoquants that are spaced or sloped slightly differently. This may trigger an adjustment in the way in which firms grow their factor inputs in the future, without directly affecting current productivity or production possibilities.

Second, for the purpose of analyzing the response of the model economy to a shock, we may want to keep the shock as neutral as possible with regards to its direct effects on business cycle indicators such as output and unemployment, so that it is not the shock itself driving the relevant fluctuations of important variables, but rather the endogenous response to the shock. This way, we will be able to focus on the ability of the economic system to generate business-cycle-like patterns rather than the characteristics of a shock required to match the data, such as its persistence. This approach is distinctly different from the one taken in the early real business cycle literature, where the properties of the TFP shock to a large extent determined the path of output, with the focus being mostly on the co-movement of other endogenous variables.

Suppose that the production function at the project level is initially the same as the one we used before, $f(k,n) = Dk^{\alpha}n^{\beta}$. At the time when factor inputs have reached the level $(k,n) = (\bar{k},\bar{n})$, it becomes evident that going forward, additional capital will be slightly more or less productive than the existing stock, where the relative change is given by $\frac{\sigma^{[k]}}{\alpha}$. Similarly, the productivity of labour beyond current employment changes by $\frac{\sigma^{[n]}}{\beta}$. The resulting production function for $k \geq \bar{k}$ and $n \geq \bar{n}$ is

$$\begin{aligned} f(k,n) =& D\left(k + (k-\bar{k})\frac{\sigma^{[k]}}{\alpha}\right)^{\alpha}\left(n + (n-\bar{n})\frac{\sigma^{[n]}}{\beta}\right)^{\beta} \\ \doteq& D\left(1 + \sigma^{[k]}(1-\frac{\bar{k}}{k}) + \sigma^{[n]}(1-\frac{\bar{n}}{n})\right)k^{\alpha}n^{\beta}. \end{aligned} \tag{1}$$

The shocks to the productivity of individual factors have a long-term impact on total factor productivity that is equal to the size of the shocks. The immediate effect on average productivity for $(k,n) = (\bar{k},\bar{n})$ is zero, however, as the productivity change only applies to additional factor inputs. The marginal products, on the other hand, are affected as soon as the shock hits.

$$
\begin{aligned}
f_k \doteq & \alpha D\left(1+\sigma^{[k]}(1-\frac{\bar{k}}{k})+\sigma^{[n]}(1-\frac{\bar{n}}{n})+\frac{\sigma^{[k]}}{\alpha}\frac{\bar{k}}{k}\right)k^{\alpha-1}n^{\beta} \\
f_n \doteq & \beta D\left(1+\sigma^{[k]}(1-\frac{\bar{k}}{k})+\sigma^{[n]}(1-\frac{\bar{n}}{n})+\frac{\sigma^{[n]}}{\beta}\frac{\bar{n}}{n}\right)k^{\alpha}n^{\beta-1}
\end{aligned}
$$

The shock $\sigma^{[kn]}$ we will consider in our simulations is a combination of the capital and labour productivity shocks of the same magnitude, $\sigma^{[kn]} = \sigma^{[k]} = \sigma^{[n]}$. As this shock has a similar effect on productivity as a TFP shock, but only to the extent that additional factor inputs are used, we will refer to it as a marginal TFP shock.

With regards to the specification of the GAC function ψ, we will follow the assumptions of proposition I-2 (2) for $a = 1$.

$$
\psi(\gamma^{[k]}, \gamma^{[n]}) = \psi(\gamma^{[k]}) = \left(\frac{\gamma^{[k]}}{\bar{\gamma}}\right)^{\chi}
$$

Only the adjustment of capital is constrained, while the labour input can be varied freely. Moreover, we will be interested in the limit for $\chi \to \infty$, where the functional form of ψ guarantees that firms choose $\gamma^{[k]} = \bar{\gamma}$ and that $\psi = 0$. This particular choice of the GAC function keeps the model as simple as possible.[2]

Calculating aggregate output from project-level output (1), using the definition of the two types of shocks, we arrive at

$$
Y \doteq D\left(1+\sigma^{[D]}+\sigma^{[kn]}(2-\frac{\bar{k}}{k}-\frac{\bar{n}}{k}\frac{K}{N})\right)k^{\alpha+\beta-1}K^{1-\beta}N^{\beta}. \tag{2}
$$

Labour input is not restricted by GACs so that, according to equation (I-4), wages will reflect the marginal product of labour as given by the atemporal production function f. Capital income can then be derived as the residual, by eliminating wages from the free entry

[2]Of course, given that the rate of TFP growth is always constant in the case where only one production factor is subject to GACs, any specification of $\psi(\gamma^{[k]})$ consistent with Definition I-1 for $a = 1$ would yield equivalent outcomes, albeit for a constant $\psi > 0$.

condition.

$$r+\delta \doteq (1-\beta)D\left(1+\sigma^{[D]}+\sigma^{[kn]}(2-\frac{\bar{k}}{k}-(1+\frac{1}{1-\beta})\frac{\bar{n}}{k}\frac{K}{N})\right)k^{\alpha+\beta-1}K^{-\beta}N^{\beta} \tag{3}$$

$$w \doteq \beta D\left(1+\sigma^{[D]}+\sigma^{[kn]}(2-\frac{\bar{k}}{k}-(1-\frac{1}{\beta})\frac{\bar{n}}{k}\frac{K}{N})\right)k^{\alpha+\beta-1}K^{1-\beta}N^{\beta-1} \tag{4}$$

2.2 The Labour Market

In contrast to the model presented in the growth section I-2, we will now allow for unemployment resulting from search frictions. Let L be the labour force, which grows at the constant rate ℓ and is assumed to be proportional to population. For aggregate employment N, the unemployment rate u is then defined by the relationship $N=(1-u)L$.

Unemployed workers find jobs at the exogenous rate ϕ.

Unemployment exists for three reasons. First, workers that are new to the labour market need to find a job. Second, there are idiosyncratic separations that happen at the exogenous rate λ. Third, whenever a project is terminated, all workers previously assigned to it become unemployed. Project terminations happen at the endogenously determined rate τ. The transition function for employment can then be written as

$$\begin{aligned} N_t &= N_{t-1}+\phi(L_{t-1}-N_{t-1})-\lambda N_{t-1}-\tau_t(1-\lambda)N_{t-1} \\ &\doteq \phi(L_{t-1}-N_{t-1})+(1-\lambda-\tau_t)N_{t-1}. \end{aligned} \tag{5}$$

2.3 The Household Sector

The household sector is represented by a continuum of infinitely lived households, whose mass in period t is equal to their inelastic labour supply L_t. Actual employment $N_t = (1-u_t)L_t$ is lower than that if the unemployment rate u_t is positive. While any risk to income or wealth is perfectly pooled within the household sector, each individual household cares only about its own consumption. A unit-mass of households alive in period t maximizes its expected utility from consumption

$$\sum_{s=0}^{\infty} \left(\frac{1}{1+\rho}\right)^{s} U\left(\frac{C_{t+s}}{L_{t+s}}\right), \tag{6}$$

where C_s is aggregate consumption in period s, ρ is the households' discount rate and U is the instantaneous utility function, which is of the CRRA variety with relative risk aversion $\theta > 0$ so that $U'(c) = c^{-\theta}$.

For a wage rate w_s in period s, the sequence of transition functions for aggregate financial wealth W_s is

$$W_{s+1} = (1 + r_s)(W_s + w_s N_s - C_s) \tag{7}$$

for all periods $s \geq t$. These transition functions, scaled to the appropriate unit size, determine the households' period budget constraints.

The households' optimal consumption pattern for periods t and $t+1$ is readily derived from (6) and (7) and takes the familiar form

$$U'\left(\frac{C_t}{L_t}\right) = \frac{1 + r_t}{(1+\rho)(1+\ell)} U'\left(\frac{C_{t+1}}{L_{t+1}}\right)$$

or

$$r_t \doteq \theta\gamma_{t+1}^{[C]} + \rho + (1-\theta)\ell, \tag{8}$$

where $\gamma^{[C]}$ is the growth rate of aggregate consumption C.

2.4 A Simple Model of Cycles: Shocks and Unemployment

This subsection introduces a basic version of the business cycle model. While it does not yet explain prolonged recessions and fluctuations in the value of capital, it does include all the mechanics that contribute to the medium-term co-movement of the core variables.

The assumption that makes this version of the model relatively simple is that employed labour can move between projects and firms. This ensures that all projects are using the same factor intensities and pay the same wages within a period.

2.4.1 Equilibrium

The perfect foresight equilibrium in our model economy is given by the path of aggregate consumption C_t that solves the households' optimization problem and is therefore consistent with the Euler equation (8) and the households' transversality condition, such that

- the interest rate is consistent with optimal firm behaviour and free entry, (3),
- the aggregate capital stock evolves according to $K_{t+1} = (1 - \delta)K_t + I_t$,
- employment changes are given by (5) while the labour force grows at the rate ℓ,
- project-level capital k grows at the rate $\bar{\gamma}$ and
- the goods market clears, $Y_t = C_t + I_t$, where output Y_t is given by (2),

given the initial values of N, L, K and k.

The only thing we have not specified and derived yet is the exit rate τ_t, which plays an important role in the employment transition function (5). Since we assume the capital market to always clear, the number of projects can be easily calculated as the ratio of aggregate capital to project level capital k_t. The number of projects thus grows at the rate $\bar{\gamma} - \gamma^{[K]}$. A negative net rate of growth of this number is one source of project terminations. We allow for the possibility that – irrespective of the state of the economy – a fixed share $\bar{\tau}$ of randomly selected projects is terminated every period due to idiosyncratic shocks. This creates turnover that can contribute to a higher unemployment rate in steady state, similar to the separation parameter λ. The third component of the termination rate is due to the assumption that projects are sector specific.

Suppose a share $\bar{\tau}$ of projects in each sector has already been shut down. With the number of firms in the consumption and investment sector in period t given by $\frac{C_t}{Y_t}\frac{K_t}{k_t}$ and

$\frac{I_t}{Y_t}\frac{K_t}{k_t}$, respectively, we can calculate the rate of project terminations τ_t^+ that are necessary given the overall growth in project numbers in each sector as

$$\begin{aligned}
\tau_t^+ &= \frac{\max\{\frac{C_{t-1}}{Y_{t-t}}\frac{K_{t-1}}{k_{t-1}}(1-\bar{\tau}) - \frac{C_t}{Y_t}\frac{K_t}{k_t}, 0\} + \max\{\frac{I_{t-1}}{Y_{t-t}}\frac{K_{t-1}}{k_{t-1}}(1-\bar{\tau}) - \frac{I_t}{Y_t}\frac{K_t}{k_t}, 0\}}{\frac{K_{t-1}}{k_{t-1}}(1-\bar{\tau})} \\
&= \max\left\{\frac{C_{t-1}}{Y_{t-t}} - \frac{C_t}{Y_t}\frac{K_t}{k_t}\frac{k_{t-1}}{K_{t-1}}\frac{1}{1-\bar{\tau}}, 0\right\} + \max\left\{\frac{I_{t-1}}{Y_{t-t}} - \frac{I_t}{Y_t}\frac{K_t}{k_t}\frac{k_{t-1}}{K_{t-1}}\frac{1}{1-\bar{\tau}}, 0\right\} \\
&\doteq \frac{C_{t-1}}{Y_{t-t}}\max\left\{\gamma^{[Y]} - \gamma^{[C]} + \bar{\gamma} - \gamma^{[K]} - \bar{\tau}, 0\right\} + \frac{I_{t-1}}{Y_{t-t}}\max\left\{\gamma^{[Y]} - \gamma^{[I]} + \bar{\gamma} - \gamma^{[K]} - \bar{\tau}, 0\right\}.
\end{aligned}$$

The total rate of project terminations is then $\tau_t \doteq \tau_t^+ + \bar{\tau}$.

2.4.2 Cyclical Properties

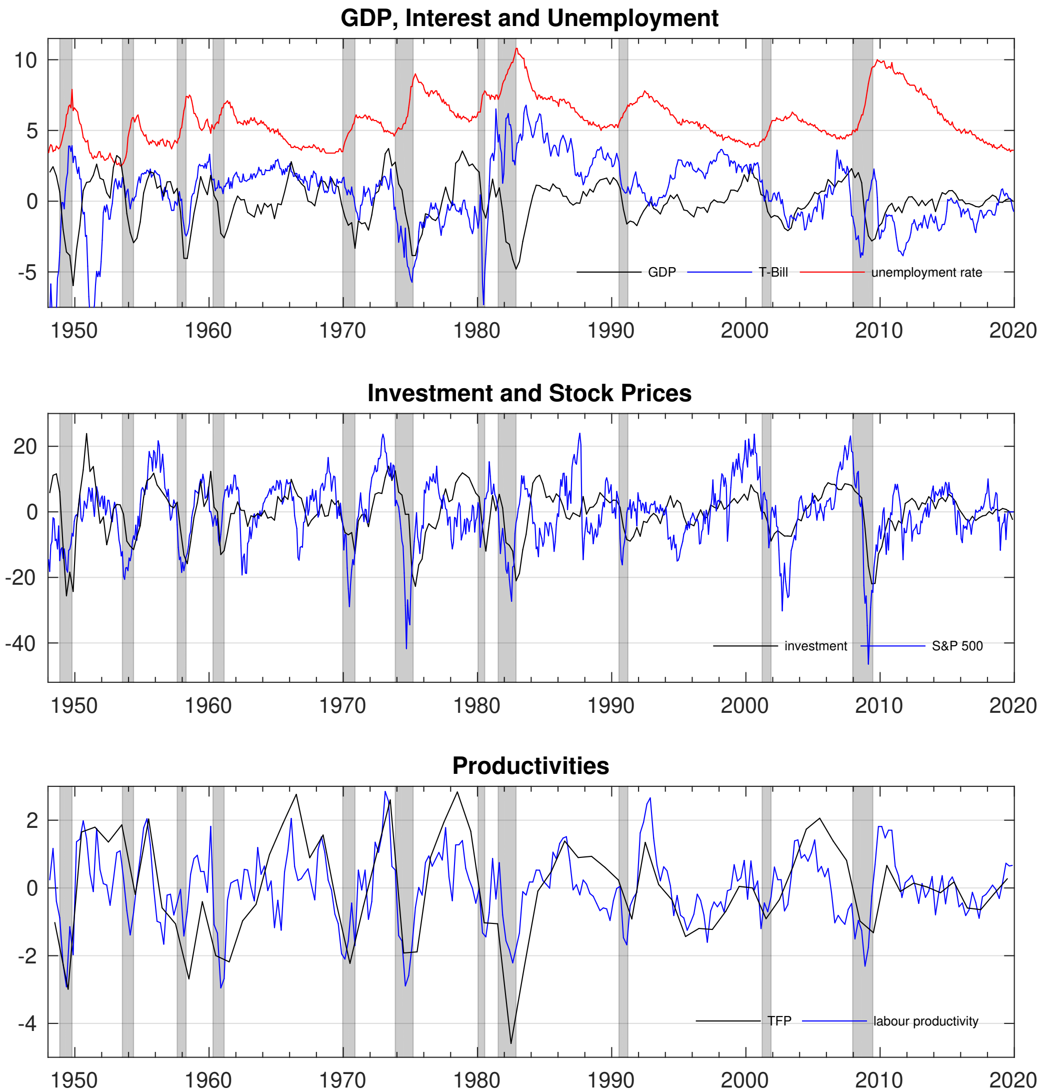
We now investigate how the model economy responds to shocks, with a particular focus on its ability to recreate patterns associated with recessions and recoveries from recessions. Rather than attempting to solve for the dynamics analytically, I will explain the mechanics of adjustments following exogenous shocks and then show simulation results obtained from a calibrated version of the model and compare them to US data.

Figure 1 depicts the time series of a number of variables created from detrended US data that have pronounced cyclical components: output and investment; unemployment; stock prices and a measure of the real interest rate; labour productivity and TFP.[3] NBER recessions are marked by vertical bars. Figure 2 highlights the patterns around recessions, showing the fluctuations of some of these variables relative to the peak of the cycle by averaging over several recessions.

The patterns are familiar. The onset of a recession coincides with the start of marked decline in output and investment, as well as a quick rise in unemployment. Stock prices drop sharply near the peak of the cycle, but usually start recovering while the economy is still

[3] The details about the detrending methods are listed in the notes to the figures for each series. The preferred choice is to use a Baxter-King high-pass filter whenever truncation of the time series is not a problem. I use the notation $\text{BK}_{\bar{T}}^{K}$ for the high-pass filter with cut-off period $\bar{T}$ (i.e. critical frequency $\omega = 2\pi\bar{T}^{-1}$) and a maximum lag length of K (see Baxter and King (1999)). HP_λ refers to the Hodrick-Prescott filter with smoothing parameter λ (see Hodrick and Prescott (1997)).

Figure 1: Time Series: US



Notes: All ordinate values in percent. Monthly data on real annualized T-Bill returns (nominal returns minus year-over-year CPI inflation) and unemployment as rates, all other data as relative deviations from trend. Quarterly GDP, investment and labour productivity data in logs detrended with HP_{1600} filter. Monthly real (CPI-adjusted) log S&P 500 data filtered with BK^{36}_{96}, annual log TFP data with HP_{100}. NBER recessions are marked in grey.

Data: Real GDP and investment from U.S. Bureau of Economic Analysis (BEA) national accounts. Unemployment rate, labour productivity, TFP (multifactor productivity) and CPI (all urban consumers, city averages) from U.S. Bureau of Labor Statistics (BLS) via FRED. T-Bill: 3-Month Treasury Bill Secondary Market Rate from Board of Governors of the Federal Reserve System via FRED. S&P 500 monthly data (closing value, first of the month) via MacroTrends and Yahoo Finance.

in recession.[4] Labour productivity and TFP drop during the recession, but quantitatively this change is much smaller than the reduction in output. Both productivity indicators start rising before the end of the recession, with labour productivity increasing faster than TFP and surpassing its pre-recession level as the economy passes through the trough. The interest rate does not follow as clear a pattern as the other variables. Highly liquid assets such as treasury bills tend to pay very low interest during and after recessions. While bond yields often drop at the beginning of a recession, this is not always the case, and on average bond yields tend to be somewhat higher after a recession than before.

The recovery of output, investment and unemployment is very slow, taking several years. Moreover, the adjustment of these variables during the recovery period is remarkably smooth and predictable. A much more comprehensive analysis of the business cycle patterns in aggregate data can be found in Stock and Watson (1999).

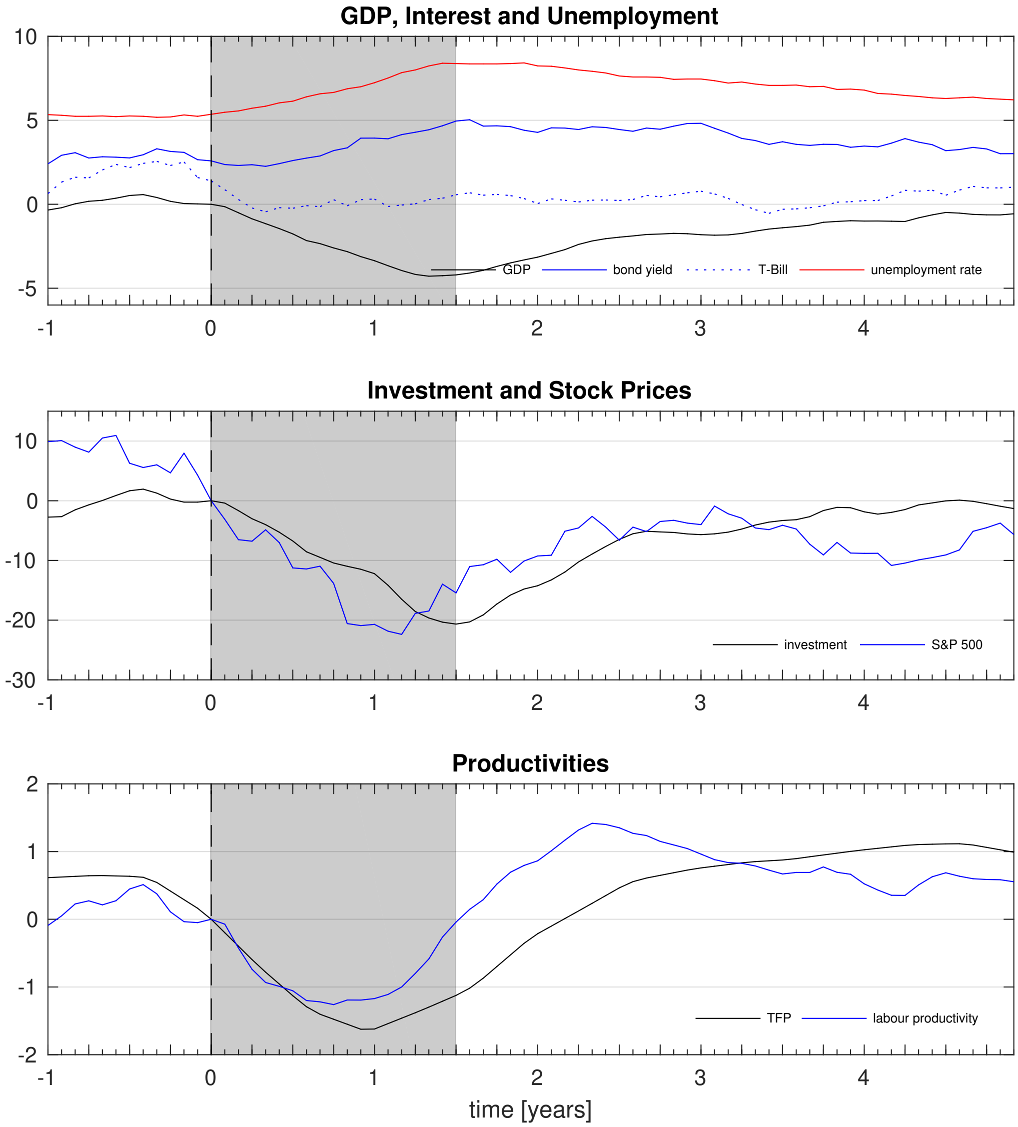
Figure 3 shows the simulated response of the economy to a one-time, unexpected negative TFP shock $\sigma^{[D]} = -3\%$ for the calibration given in table 1 in a perfect-foresight setting.[5] The top two panels present variables that are comparable to the data presented in figure 2. Qualitatively, the patterns match for the recovery period. Note that both figures are aligned at period $t = 0$ when the downturn begins. While the recession itself lasts for several quarters to a year in the data, figures 1 and 2, the move to the trough of the cycle is instantaneous in this version of the model, figure 3. The reason for this and how this aspect can be made more realistic will be discussed in subsection 2.5 below.

The initial drop in simulated output of about 6% is twice as large as the productivity change directly caused by the shock. Half of the output gap is explained by a rise in unemployment from 5% to about 8%, consistent with Okun's law. After the shock, output recovers very slowly and unemployment remains elevated for several years, dropping at a rate that is much smaller than what should be expected considering that individual unem-

[4] In the 1981 and 1990 recessions, the S&P 500 had almost fully recovered by the time the economy reached the trough of the cycle.

[5] Unless stated otherwise, all simulations are conducted at monthly frequency.

Figure 2: Average Time Series Recession Patterns: US



Note: Average of the 1973, 1981, 1990, 2001 and 2007 recessions, lined up at NBER peak month ($t = 0$). All deviations from trend normalized to zero at $t = 0$.
Data: Bond yield is Moody's Seasoned Aaa Corporate Bond Yield via FRED. All other series are constructed and filtered as in figure 1.

Table 1: Calibration

variable	interpretation	value
θ	relative risk aversion	2
ρ	discount rate	0.02
ℓ	labour force growth	0
α	capital coefficient	$\frac{2}{3}$
β	labour coefficient	$\frac{2}{3}$
D	productivity constant	1
δ	depreciation rate	0.06
$\bar{\gamma}$	project-level capital growth	0.04
ϕ	job finding rate	2
λ	exogenous job destruction rate	$\frac{3}{4}\frac{0.05}{0.95} \approx 0.039$
$\bar{\tau}$	project destruction rate	$\frac{1}{4}\frac{0.05}{0.95} \approx 0.013$
	persistence of TFP shock	0.5
	persistence of marginal TFP shock	0.9

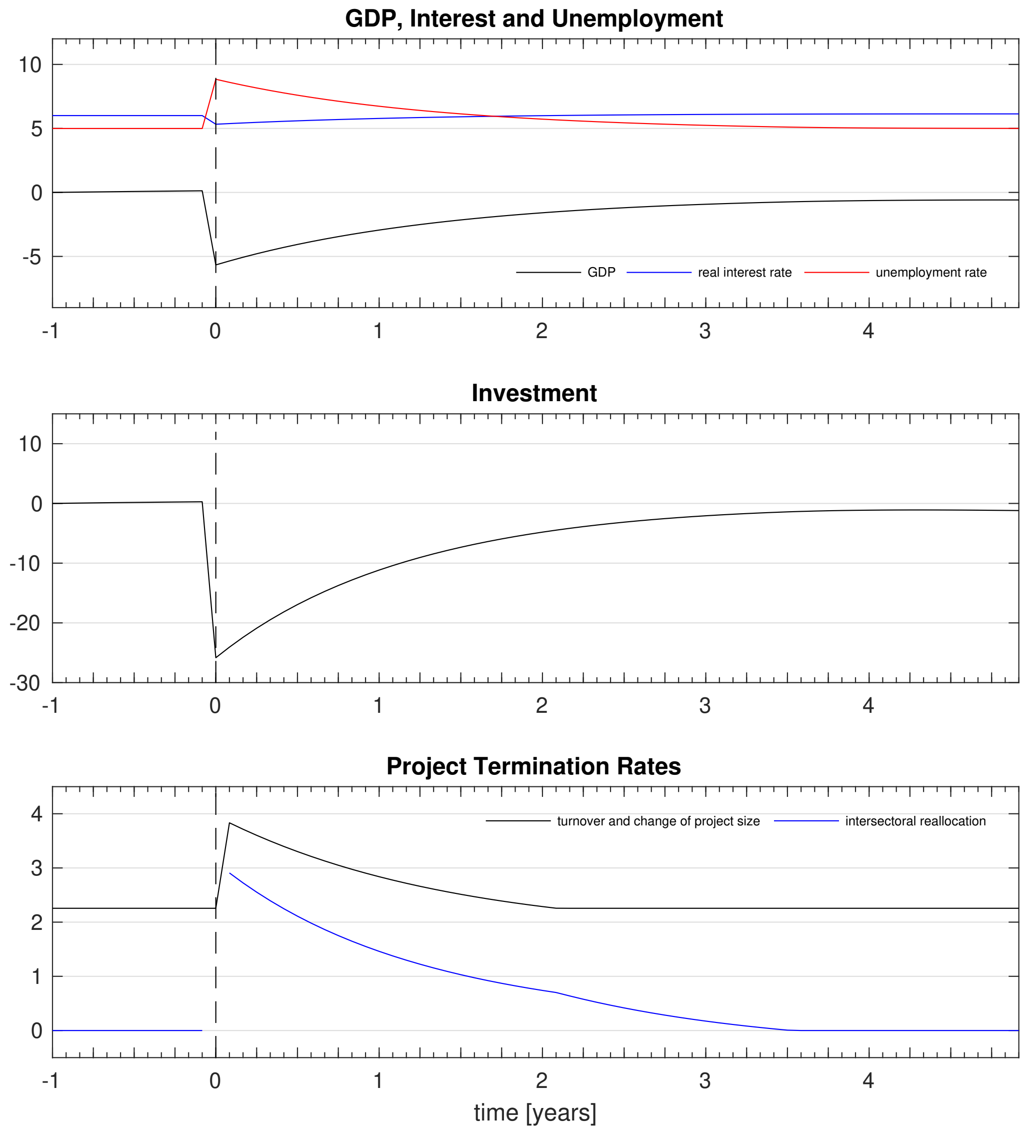
All rates, including shock persistence, are annualized.

ployment spells last two quarters on average. Two things need to be explained: The depth of the recession, i.e. the initial output gap and unemployment level, and the persistence of the deviation of these variables from their long-term levels.

Fundamentally, the reallocation process triggered by the negative TFP shock is exactly the one familiar from the neoclassical growth model. The initial reduction in current income, exacerbated by the drop in the marginal product of capital and thus the interest rate, prompts the household to reduce savings, necessitating a reduction in investment. The resulting slow reduction of the capital-labour ratio over the following periods compared to the previous trend then makes the interest rate rise again over time, stabilizing the economy even if the shock is permanent. In our setting, this adjustment process is modified by the interaction of project dynamics with labour market frictions.

Due to the sector-specific nature of projects, the reallocation of resources away from the investment sector triggers additional project terminations and exit, leading to the sharp rise in unemployment and an even bigger drop in output. While the capital associated with these terminated projects becomes available for use elsewhere, the corresponding labour cannot

Figure 3: Simple Model: TFP Shock $\sigma^{[D]} = -3\%$



Note: Simulated impulse response. All ordinate values in percent. All rates annualized. Shock applied at time $t = 0$. GDP and investment as logs detrended with BK_{96}^{96}.

immediately be reemployed due to frictions. The consequence is that the economy-wide factor input ratio changes, which, together with the income drop, sets in motion a feedback loop that amplifies the effect of the original investment-reducing shock: The initial drop in employment both reduces output, creating a negative transitory income shock to households, and raises the capital-labour ratio, thus lowering the interest rate even more. Both of theses changes prompt households to shrink their saving and with it investment further. This effect strengthens the response of investment, the interest rate, unemployment and output to the negative TFP shock, making the recession deeper.

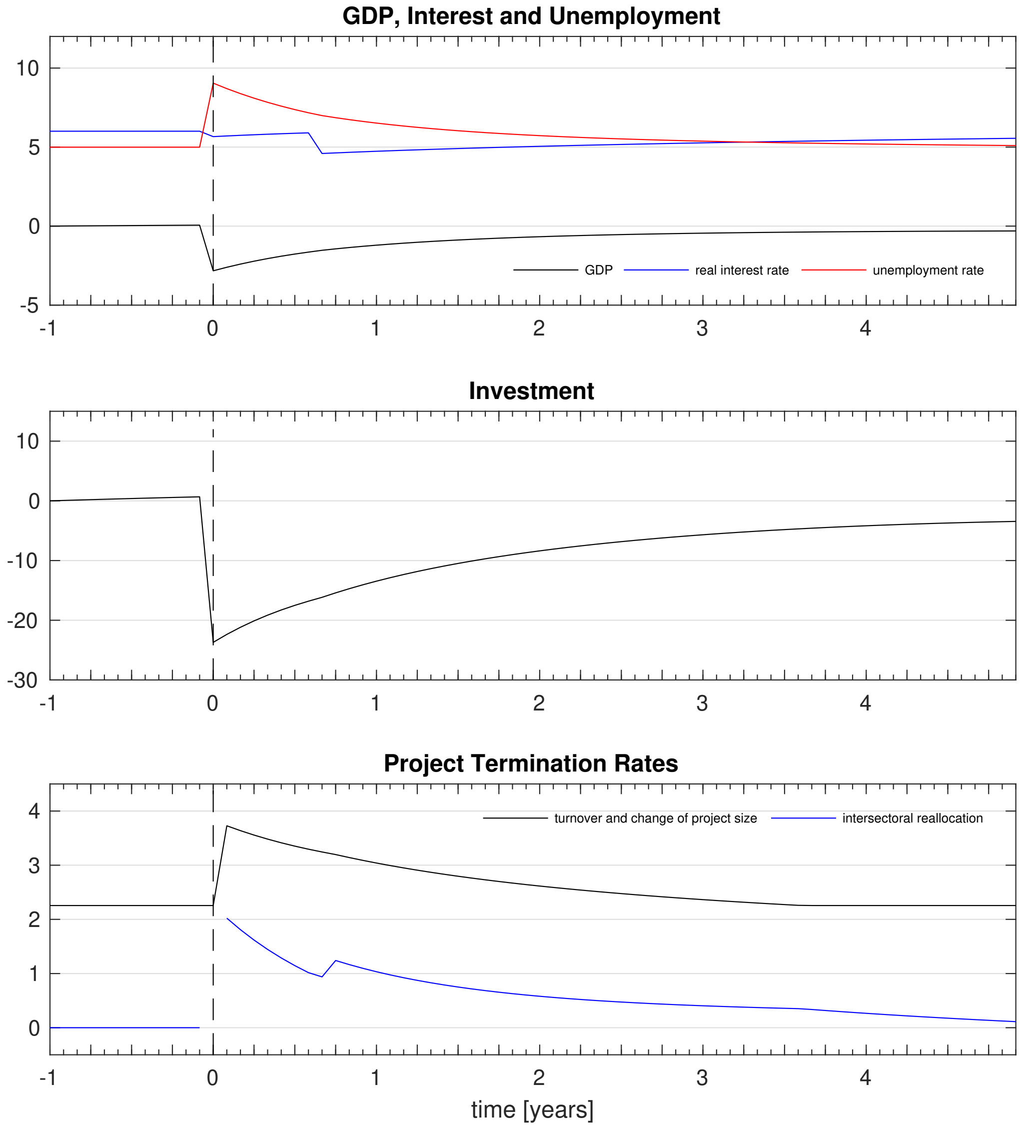
The reason why unemployment remains high after the initial response to the shock is ongoing turnover in the firm sector. There are two aspects to this. First, in order to implement the reduction of the aggregate capital stock below its trend path, which is the result of the lower investment, there needs to be an increase in the *net* termination rate, because the project-level capital stock follows a fixed trend. Second, as investment rises again towards its trend level, more intersectoral reallocation is required, which leads to additional project terminations in the shrinking sector. The bottom panel of figure 3 illustrates these two sources of economic turbulence. The black line shows how net project exit increases and remains elevated following the shock, whereas the blue line tracks the effect of excess project shutdowns in the shrinking sector.[6]

This higher turbulence resulting from ongoing exit of projects leads to higher unemployment and thus lower output, with the effect being related to how relative sector sizes deviate from their normal level and how they change.[7] Following investment-reducing shocks, both effects work in the same direction and contribute to dynamics that create the impression of a slow recovery from the shock.

[6]The observation that project exit is associated with recessions as well as the notion that capital is being repurposed elsewhere relates our model to a literature that explains patterns regarding the reorganization of production, the implementation of innovations and the replacement of physical capital over the business cycle. See for example Cooper and Haltiwanger (1993), Caballero and Hammour (1994), and Cooper, Haltiwanger, and Power (1999).

[7]These effects, being tied to changes in sectoral composition, result in similar patterns as mismatch unemployment. See Şahin et al. (2014).

Figure 4: Simple Model: Marginal TFP Shock $\sigma^{[kn]} = 4\%$



Note: See figure 3 for details.

A negative TFP shock can explain recession-like events in our model. However, it does so by inducing a relatively large direct effect on output. Almost half of the roughly 6% drop in GDP we see in figure 3 is immediately explained by the productivity effect of the shock. Moreover, while the rise in unemployment of almost 4 percentage points in our simulation is consistent with what happens in the average recession shown in figure 2, both the size of the TFP shock and the extent of the output response are about 50% bigger than what is suggested by the data.

Figure 4 shows the effect of a marginal TFP shock $\sigma^{[kn]} = 4\%$, which, in contrast to the TFP shock, does not have any immediate productivity impact at all. Instead, it works by triggering a reallocation towards more labour intensive production.[8] The response of the economy to this type of shock is very similar to what happens following the TFP shock. The main difference is that the output reduction is more moderate, thanks to the absence of the direct productivity effect.[9] This shows that it does not take changes to average productivity to trigger recession events in our model, and that to an extent the co-movement of variables following a recession can be independent of the exact nature of the shock.

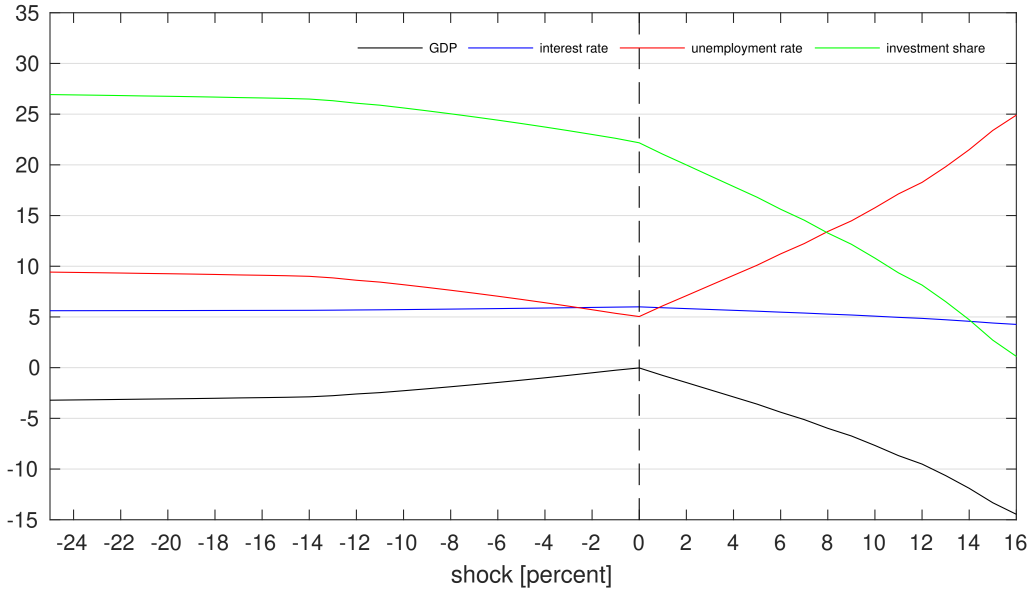
We have so far focused on shocks that cause a reduction in investment, which is one of the typical features of recessions. As we do not observe dramatic increases in unemployment associated with increases in the investment share, it is natural to ask if there is an asymmetry between the response to investment increasing and investment reducing shocks.

Such an asymmetry is very clearly visible in figure 5, which shows the initial response of various variables of interest to marginal TFP shocks of different sign and magnitude. The

[8]Higher productivity beyond the current scale of production means higher marginal products, and the income of the freely adjustable factor thus rises. Given that average productivity does not change immediately, the residual return r on the factor that is subject to GACs falls. To restore a rate of return on capital that is consistent with consumption growing at the same rate as output, individual projects must become more labour intensive. This requires the economy-wide capital stock to fall.

[9]Another difference compared to the TFP scenario is the jump in the interest rate and one of the termination rates that occurs several months after the shock arrives. The reason for this lies in the nature of the technology shock. The productivity change applies to labour input beyond the level reached at the time of shock. As a consequence of the recession triggered by the shock, unemployment rises and labour intensity initially drops throughout the economy, so that the higher marginal products are not immediately relevant. It takes several months for the labour intensity to reach its pre-shock level, at which point the marginal product of labour increases and the interest rate drops.

Figure 5: Asymmetry of Cycles: Initial Response as Function of Shock



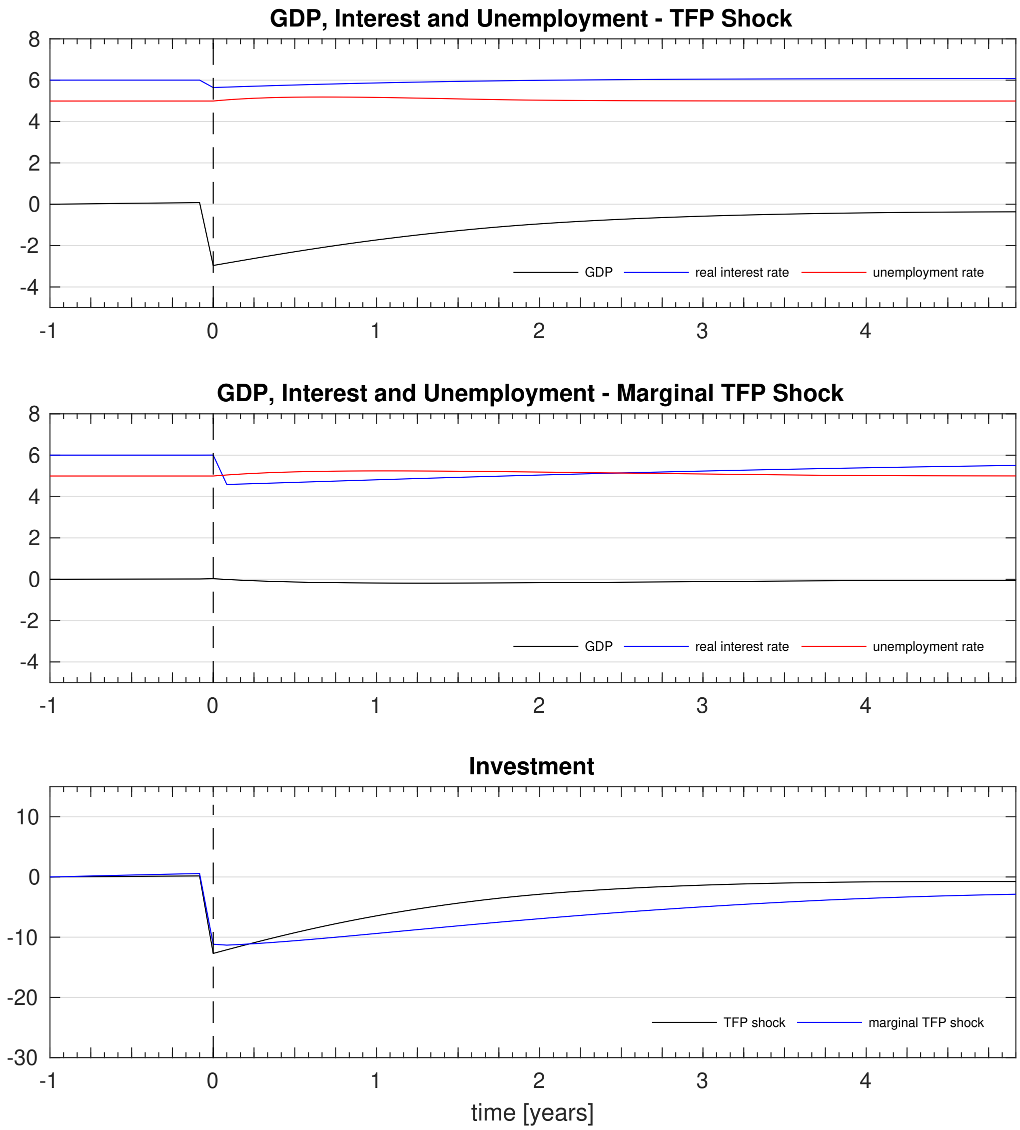
Note: All ordinate values in percent. Interest rate annualized, GDP as relative deviation from steady state level. The shock is a marginal TFP shock $\sigma^{[kn]} \in [-25\%, 16\%]$.

value of the shock is shown on the horizontal axis, ranging form -25% to 16%. The $\sigma^{[kn]} = 0$ point, marked by a vertical dashed line, corresponds to steady state values where no shock occurs. While a large investment-reducing shock $\sigma^{[kn]} = 10\%$ raises the unemployment rate by 10.7 percentage points to almost 16%, an opposite sign shock of the same magnitude increases u by only 3.2%.

The same effect that amplifies investment-reducing shocks dampens investment-increasing ones. Suppose a shock arrives that would make households want to reduce consumption and increase their savings by a certain amount in the absence of frictions. The reallocation of productive resources from the consumption to the investment sector increases unemployment, reducing income and the interest rate and therefore the incentive to save and invest. Figure 5 shows how both the dampening and amplifying effects get stronger with the magnitude of the shock. This phenomenon can help explain the absence of “anti-recessions” in macroeconomic data.

Finally, figure 6 highlights the importance of sector specific production for the response

Figure 6: Simple Model, No Sector Specific Production: $\sigma^{[D]} = -3\%$, $\sigma^{[kn]} = 4\%$



Note: Simulated impulse response. All ordinate values in percent. All rates annualized. Shock applied at time $t = 0$. GDP and investment as logs detrended with BK_{96}^{96}.

to shocks and the ability of the model to create recession-like events. It shows the response to the same TFP and marginal TFP shocks as in figures 3 and 4, respectively, but without the sector-specific nature of projects. If it is possible to switch between creating consumption and investment goods, the unemployment response to shocks almost disappears. In the case of the TFP shock, the 3% drop in output is simply due to the direct effect of the shock on productivity. The adjustment process shown is largely what one would see in the standard neoclassical growth model. Beyond the obvious unemployment effect that is directly caused by the imposed friction, comparing figures 3, 4 and 6 clearly shows the extent of the amplification of the adjustment process caused by the project terminations resulting from intersectoral reallocation. The drop of investment is about twice as large in figures 3 and 4 compared to figure 6, and similarly the immediate reduction of the interest rate following the TFP shock[10] is almost twice as strong with sector specific projects.[11]

2.5 The Full Model: Recessions

We will now develop the full version of the business cycle model. Compared to the simpler model of the previous section 2.4, we will gain two important features: longer lasting recessions, during which unemployment rises and output falls, and changes in asset prices, most importantly stark asset price drops at the beginning of recessions. The only change necessary compared to the simple model presented above is to drop the assumption that firms can hire already employed labour from competitors. From now on, all new hires will have passed through the labour marked and come out of a period of job search.

In a frictionless RBC model, or in the model without sector specific production shown in figure 6, moving production to a different sector is always possible at an opportunity cost

[10] The change is -0.67 percentage points with reallocation frictions vs. -0.36 percentage points without. As the interest rate effect is partly delayed in the case of the marginal TFP shock, a direct comparison is not as straightforward in this scenario.

[11] The frictions preventing instantaneous intersectoral reallocation of labour connect the mechanisms presented here to an older literature on the importance of sectoral shifts for aggregate unemployment. Lilien (1982) finds evidence that the majority of the cyclical component of unemployment in the 1970s was due to intersectoral shifts. At the time, the main focus was on the question of whether such patterns can be explained by aggregate shocks, see Abraham and Katz (1986) and Hosios (1994).

of one. Increasing the consumption share by one percentage point can be accomplished by reducing investment by one percentage point. Even in the simple business cycle model of section 2.4, relatively large adjustments the components of GDP were possible, albeit at a higher opportunity cost. If households wanted to increase consumption by 1% of GDP, resources could be freed up by shutting down investment projects. The labour employed in these projects would become unemployed and thus unavailable, but the capital could be moved to consumption goods production. After reallocating some labour between projects to equalize labour intensities, the consumption sector would end up with substantially more capital but only slightly more labour. Due to this change in production technique, it would take more than a one percentage point reduction in investment to increase consumption by one percentage point.

Shutting down between-firm hiring limits the amount of labour available for shifting production to the current flow out of unemployment, ϕuL. It is still possible to free up a large amount of capital in one sector to put it to use in new projects in another, but all these new projects have to share the limited labour available for hire. This effectively constrains the rate at which production can be moved from one sector to another. If the supply of goods cannot respond quickly enough to changing demand, prices must adjust. We should therefore expect the price of investment goods to change relative to that of consumption goods in the case of strong shifts in relative demand between sectors.

2.5.1 Consumption and Saving

Reconsider the households' problem (6) and (7) for the more general case where the price of consumption $p^{[C]}$ or the price of investment $p^{[I]}$ may deviate from unity. In this case, we can write the budget constraint in period s as

$$p_s^{[I]}W_{s+1} = (1 + r_s)(p_s^{[I]}W_s + w_s(1 - u_s)N_s - p_s^{[C]}C_s). \tag{9}$$

A unit of wealth W should be interpreted as a claim to a single unit of capital. The interest rate r is then the return on this wealth also measured in units of wealth (rather than consumption units), which would be the return on capital net of depreciation in equilibrium.

Defining the relative price of capital as $q = \frac{p^{[I]}}{p^{[C]}}$, the Euler equation describing optimal consumption under the objective (6) and the constraint (9) is

$$U'\left(\frac{C_t}{L_t}\right) = \frac{1+r_t}{(1+\rho)(1+\ell)} \frac{q_{t+1}}{q_t} U'\left(\frac{C_{t+1}}{L_{t+1}}\right) \Leftrightarrow$$
$$\left(\frac{C_t/q_t^{\frac{1}{\theta}}}{L_t}\right)^{-\theta} = \frac{1+r_t}{(1+\rho)(1+\ell)} \left(\frac{C_{t+1}/q_{t+1}^{\frac{1}{\theta}}}{L_{t+1}}\right)^{-\theta}. \qquad (10)$$

For a constant relative price of investment goods q, including the special case where $q = 1$, this Euler equation is identical to the one derived earlier, equation (8).

Let t be the current period, and $\left(\bar{C}_s\right)_{s=t}^{\infty}$ the sequence of consumption levels that is optimal if $q_s = 1$ for all $s \geq t$. Suppose that the price of assets q is allowed to vary by a small amount, but is known to return to one within a finite time by period T. It is easy to see that as long as the changes are small enough not to affect the households' lifetime resources in a significant way, a θ% change in q_s results in a 1% change in C_s compared to $\bar{C}_s$ for $s < T$.

If, for example, households would like to liquidate a part of their savings to increase consumption, but doing so is not immediately possible due to production constraints, a temporary drop in asset prices q dissuades them from converting assets to consumption at a high rate and makes a gradual adjustment of consumption possible. This is what happens during recessions in our model.

2.5.2 Intersectoral Reallocation

Next, consider what happens within the firm sector if the relative demand for consumption and investment goods changes.

Suppose that up to this point, both sectors have been growing at the same rate, and constraints affecting the adjustment of productive capacity haven not been binding. Production is undertaken by firms running a large number of projects that all use the same production technique and are assigned to either the consumption or investment sector. Let project-level employment at the beginning of the current period be $\underline{n}$.[12] Now, unexpectedly, the demand for consumption goods increases before any production decisions are made.

Very small demand changes can typically be accommodated without any further macroeconomic effects. If the number of projects needs to be reduced anyway, terminations will be concentrated in the sector with reduced demand. If the number of projects is increasing, relatively more entry takes place in the growing sector. The aggregate rate of project terminations and thus unemployment remains unaffected. This mechanism operates up to the point where all regular exit happens in the shrinking sector or all regular entry takes place in the growing sector.[13]

Further adjustment comes at the expense of increased unemployment. The mechanism is exactly the same as the one described in section 2.4 above. To increase output in the consumption sector, additional projects in the investment sector are terminated. Their capital becomes available to start new projects in the consumption sector, but their labour force becomes unemployed and unavailable for immediate rehire. The available flow of new employment ϕuL is shared among all projects such that each of them has the same labour force n. A larger reallocation towards consumption means that there are more new consumption projects that start out without any employees, so that the project-level employment that can be achieved with the given amount of additional labour is lower. This mechanism works as long as the new employment ϕuL is large enough to ensure that all existing and new projects can be endowed with the same amount of labour. Once project level employment

[12] This is the amount of labour still available to the firm, after idiosyncratic separations, $\underline{n} = (1 - \lambda)n_{t-1}$.

[13] The capacity to use this costless adjustment mechanism shrinks with the period length, as within-period flows get smaller. The possibility to accommodate any discrete demand changes fully disappears in the continuous time limit. The extent to which this mechanism is available depends on the project turnover parameter $\bar{\tau}$ among other factors. For arbitrarily chosen period lengths, this might be considered a meaningless artefact tied to the particular numerical solution strategy of the model. If, however, the period length aligns with a relevant response time or decision lag or frequency inside the firm sector, this ability to fully absorb small adjustments could be a relevant feature of the model.

is as low as $\underline{n}$, all new hires are absorbed by newly created projects.[14]

If consumption needs to grow beyond the point where $n = \underline{n}$, new effects arise. As more capital is reallocated to the consumption sector, more projects are started up, which share the available labour. Employment n in these new consumption projects is below $\underline{n}$. The marginal product of labour is thus high in these projects, resulting in relatively high real wages for newly hired workers. Existing firms in the consumption sector have no interest in hiring at these high wages, but since their exiting employees cannot move to other firms, there is no pressure for them to adjust their labour force or to pay a wage that deviates from the marginal product, which is lower than in new projects as employment $\underline{n}$ is larger. While new projects earn zero profits – after all there is free entry exit in both sectors – existing consumption projects, which are lucky to be endowed with more labour that is not immediately mobile, are able to earn a rent. Investment sector projects use the same production technique with employment $\underline{n}$ as existing consumption sector projects. Yet, they are not earning a rent. There is ongoing exit in the sector and profits are zero as firms are selling their output at a lower price $q < 1$.[15]

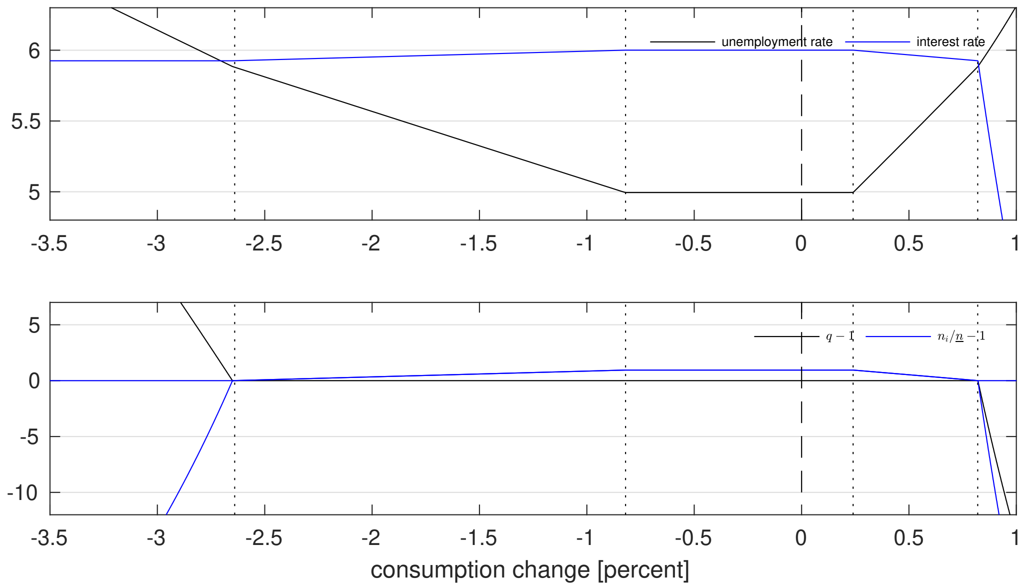
Just like for the first two simpler scenarios, we can derive the interest rate and output by sector in the $q \neq 1$ case. The interest rate is pinned down by the requirement that marginal projects in the consumption sector, which are a newly created ones in our example, earn zero profit. This means that their capital share must be sufficient to pay exactly the market return on the employed capital.

$$f(k,n) - (r + \delta)k - f_n(k,n)n = 0$$

[14]In the model of section 2.4, this was not a critical point, as for even larger reallocations, the newly created projects could directly hire workers from existing ones, so that project-level employment could continue dropping below $\underline{n}$ without any additional effects. This channel of moving labour between projects is unavailable now.

[15]Effects related to changes in the relative price of the output of different sectors have not featured prominently in the RBC literature. Kydland and Prescott (1982) discuss in some detail both q theory and a simple neoclassical setting where the opportunity cost of investment relative to consumption is always one, before dismissing both in favour of their time-to-build approach.

Figure 7: Firm Sector Outcomes as Function of the Scale of Intersectoral Reallocation



Note: All abscissa and ordinate values in percent. Dotted lines mark where regime switches occur. Simulation based on the calibration summarized in table 1, with the parameter change $\bar{\tau} = 0.75$.

Given this interest rate, the same condition for the marginal firm in the investment sector, an established firm in our example, pins down the relative price of investment goods.

$$
\begin{aligned}
qf(k, \underline{n}) - (r + \delta)k - qf_n(k, \underline{n})\underline{n} = 0 \Leftrightarrow \\
q = \frac{f(k, n) - f_n(k, n)n}{f(k, \underline{n}) - f_n(k, \underline{n})\underline{n}}
\end{aligned}
\tag{11}
$$

Finally, the rent of supramarginal projects, which are established consumption projects in our case, is determined by the extent to which they enjoy a higher output price compared to projects in the other sector using the same factor input combination.

$$
\pi = (1 - q)f(k, \underline{n})
$$

Figure 7 shows the effect of changes in consumption demand on project-level employment, the relative price of investment goods, the interest rate and unemployment. Small changes between about -0.8% and 0.2% can be accommodated without any adjustments to the interest rate or unemployment. Larger changes up to -2.7% or 0.8% lead to the effects familiar from the model in section 2.4. The interest rate drops with the amount of the reallocation, and rising unemployment due to sector-specific nature of projects implies lower labour intensities. For even larger adjustments, the bifurcation of labour intensities occurs and the relative price of investment changes.

2.5.3 Equilibrium and the Response to Shocks

For the most part, the equilibrium and dynamics of this model are identical to those of the simpler version presented in section 2.4 above. The only difference arises in situations where a sufficiently large shock triggers a desired adjustment of production patterns that cannot be accommodated immediately without creating heterogeneity in production techniques. In this case, the relative price of capital goods q changes, which makes the household tolerate a slower consumption response to the shock than what would otherwise be optimal. Such states of ongoing production shifts with $q \neq 1$ are temporary, but may persist for several periods.

Starting at a steady-state situation, consider the response to a single shock in period t that triggers a planned consumption increase and is large enough to cause q to fall below one. The initial period following the arrival of the shock sees the creation of new consumption projects with lower than average labour input and thus the split of production techniques into two groups. In the next period, another group of new consumption projects is formed as reallocation of capital from the investment to the consumption sector continues. These projects, however, can be endowed with more labour than the new projects created in the previous period, for two reasons. First, the flow of available labour increases due to the higher unemployment rate. Second the pressure to move projects from investment to consumption is lower with part of the necessary adjustment already being done. The indicator for how urgently resources need to be reallocated between sectors is the relative price of goods q. With the labour endowment n of new projects rising, while the labour input of pre-shock projects $\underline{n}$ keeps on falling at the rate λ due to worker turnover, q is bound to get closer to 1, as seen in equation (11). With these period $t+1$ projects reaching a higher labour input than period t projects did and thus paying a lower wage, those projects from the previous period raise their employment to the same level. All projects created after the arrival of the shock thus use the same production techniques in any period, and a further bifurcation in labour intensities does not occur.

After the initial split in production techniques and the drop in q, the situation slowly returns to normal over the following periods. While pre-shock firms get less labour intensive, the group of post-shock firms grows larger and more labour intensive. As production techniques converge, q approaches its steady-state value of one, which is reached in finite time, no later than when the desired adjustment is accomplished. The labour market frictions only slow down the adjustment process temporarily.

While the adjustment to a single shock will only create a split into two production techniques, it is possible that the arrival of further shocks at a time when $q \neq 1$ creates new marginal groups firms and thus further temporary divergence of labour intensities.

Note that with the use of different labour intensities for $q \neq 1$, aggregation of the whole production sector into a single Cobb-Douglas production function is, strictly speaking, not possible anymore. If we nonetheless calculate an empirical TFP measure based on the aggregate production function we observe for $q = 1$, we find that productivity is slightly below trend due to the inefficiency resulting from the use of divergent production techniques.

2.5.4 Simulation Results

Figure 8 shows the simulated response to an 8% marginal TFP shock using the same calibration as before. While the patterns are very similar to those seen previously in figure 4 for the recovery period, the striking new feature is the emergence of a recession period of 6 months, during which unemployment rises continuously while investment and output fall. At the beginning of the recession, asset prices q drop by about 10%, before recovering slowly while the economy moves from peak to trough. This drop in asset prices may seem small compared to the drop in stock prices reported in figure 2, but remember that the S&P 500 index lists highly leveraged companies. For a debt-to-equity ratio of 1.5, a 10% reduction in average asset values corresponds to a roughly 25% reduction in equity values, which is in line with the stock market declines observed in typical recessions.

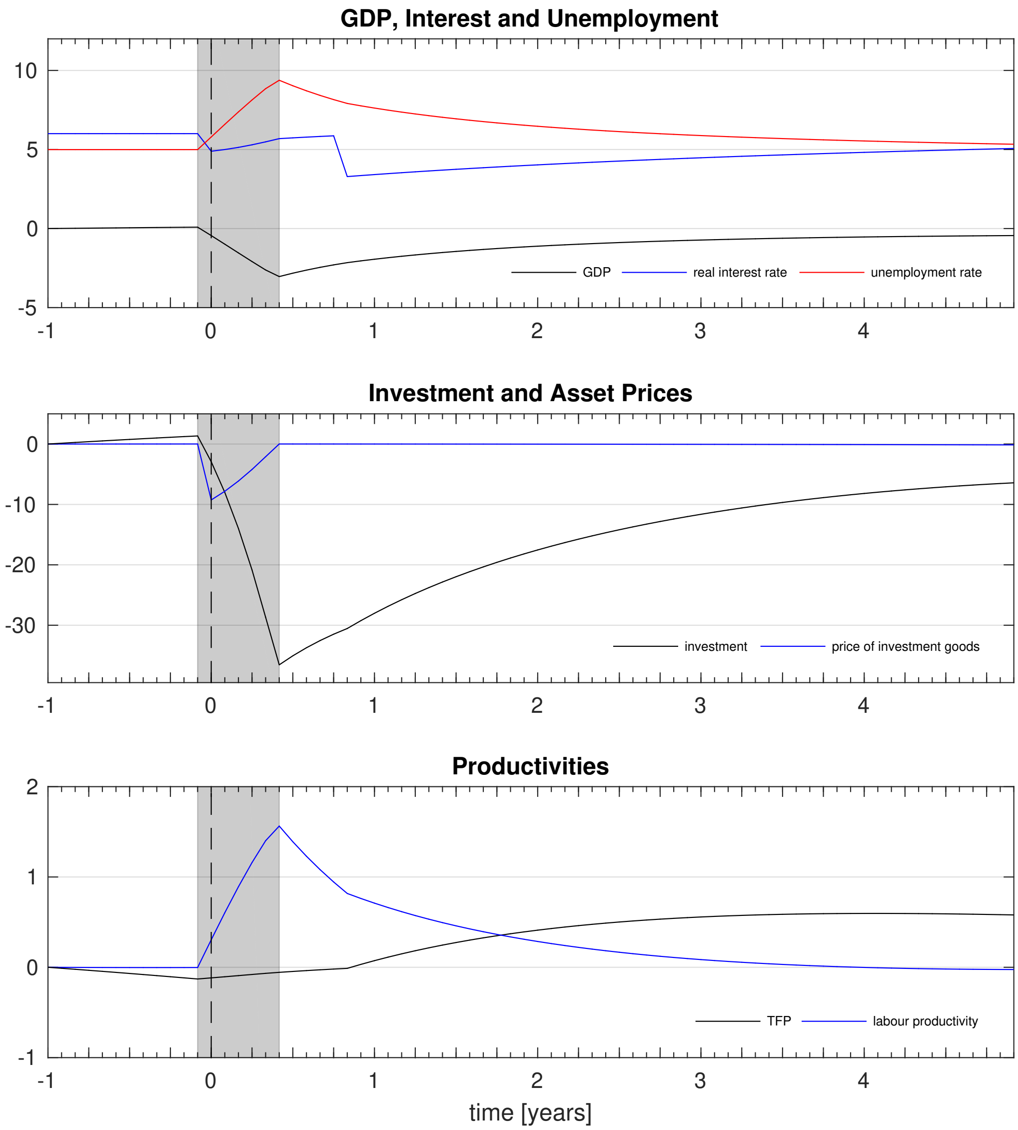
Overall, the model does a good job in replicating many of the patterns we associate with business cycles and in particular with recessions. Remember that this is really just the simple growth model from section I-2 with very basic labour market frictions added as the only new feature. Moreover, the calibration used was chosen mostly for consistency with the macroeconomic literature rather than being tuned for delivering the best quantitative results.

We will end this section on simulation results with three exercises that are designed to showcase the ability of the model to produce realistic business cycle fluctuations. None of them should be interpreted as a serious attempt to calibrate the model to the data. In all three cases, the calibration is based on that summarized in table 1, which we have used so far.[16]

The first experiment is to match a simulated quarterly time series to resemble the US data from 1948 to 2019 shown in figure 1. Starting with an economy in steady state in the initial period $t = 1$, a marginal TFP shock is chosen for each period $t = 2 \dots 288$ to best match the unemployment rate and investment share patterns in the data for the corresponding period, 1948:Q2 to 2019:Q4, proceeding period by period. Figure 9 shows the resulting simulated

[16]For the first two simulation exercises, I continue to use the perfect foresight setting, in which any shocks are technically entirely unexpected. While this approach is equivalent to the full stochastic rational expectation solution up to a linear-quadratic approximation as far as risk attitudes are concerned, there are certain aspects of expected future states that it is bound to miss. It does not account for the expected rents some firms can earn if a large shock leads to a recession. As these expected rents are extremely small, and their only effect is to encourage a larger number of projects in any state, which should not impact dynamics in an important way. Moreover, as all sufficiently large shocks increase turbulence and thus unemployment, households that do not account for the possibility of future shocks tend to underestimate future unemployment. However, as long as this effect is not dramatically state dependent, its effect on the dynamics of interest is likely to be small as well.

Figure 8: Full Model: Marginal TFP Shock $\sigma^{[kn]} = 8\%$



Note: Simulated impulse response. All ordinate values in percent. All rates annualized. Shock applied at time $t = 0$. All variables except interest rate and unemployment rate as logs detrended with BK_{96}^{96}.

series.[17]

The simulated series are able to capture a number of features of the data well, including the patterns of unemployment, investment and output. While the simulation does not replicate the large swings in the interest rate that occurred since the 1940s, especially up to 1980, it does match certain aspects of the interest rate fluctuations, such as the decline of the interest rate after 1984 or the drop and recovery of interest rates following the recessions since 1990, albeit at a much lower amplitude than the data. Overall, this simulation exercise shows that the model is capable of producing realistic co-movement of macro variables, as it can replicate the cyclical behaviour of multiple time series by varying only a single shock parameter per period.

In the second exercise, I simulate the economy at quarterly frequency under a sequence of randomly generated TFP shocks. With a probability of 25% per period, a new shock arrives, which is drawn from a normal distribution with a standard deviation of 2%. Table 2 compares the standard deviations and correlations with output of several variables between US data and the simulated economy. The data are from 1948:Q1 to 2019:Q4, the same 72 years or 288 periods shown in figure 1. The simulated series cover 288 periods, too. All the data moments reported are highly significant, with the exception of the correlation of the real interest rate with output.[18]

For the most part, the moments generated from model output are very similar to the data moments. The two variables that do not match their empirical counterparts are consumption and the real interest rate, which are both much more volatile in the data than in the simulations. This is a frequent feature of business cycle models built around a simple household sector model like ours, including many RBC models.[19,20]

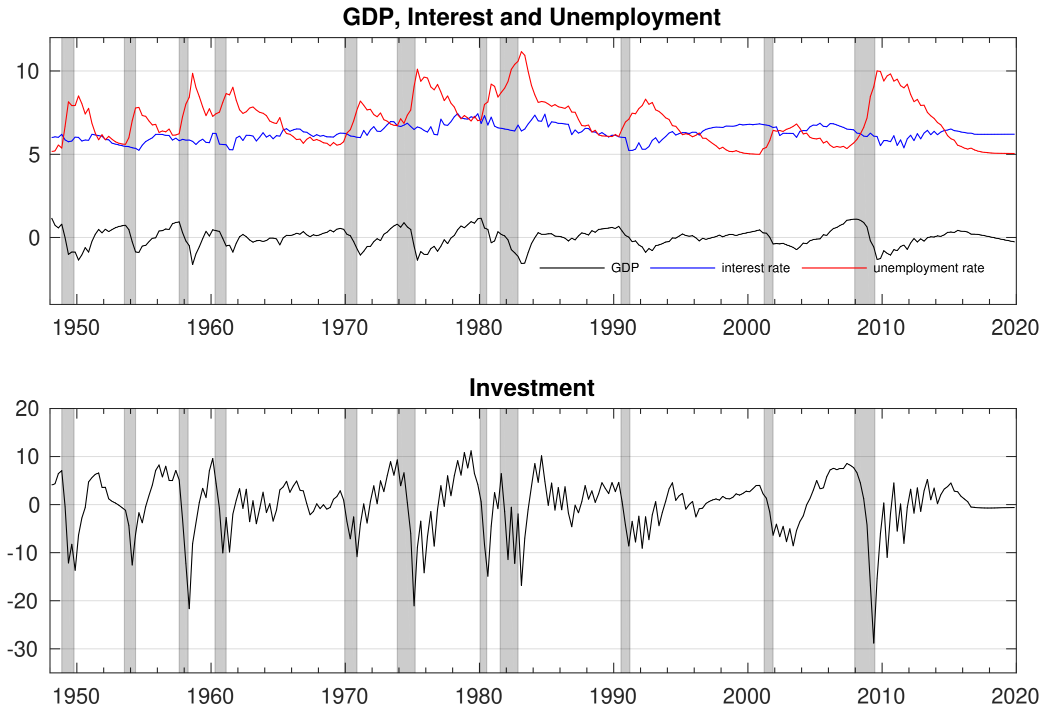
[17] Details about the matching process can be found in the notes to figure 9. Matching period by period is not a good way of obtaining an estimate of the shock process. For that purpose, a global criterion applied to a carefully calibrated model would clearly be preferable.

[18] I use a moving block bootstrap approach to obtain a variance measure for the moments that reflects their stability over time. This is done by splitting the 72-year sample into twelve overlapping 24-year blocks.

[19] See for example Rebelo (2005) for a broad assessment of the capabilities of RBC models and the impact of a variety of extensions on their ability to replicate cyclical patterns in the data accurately.

[20] In the data, the share of consumption in output remains remarkably stable over the cycle, implying

Figure 9: Matched US Time Series - Marginal TFP Shock



Note: All ordinate values in percent. Interest rate rates annualized. All variables except interest rate and unemployment rate as logs detrended with HP_{1600}.
The model is calibrated as in table 1, except for the shock, which is permanent (persistence=1). The value of the shock $\sigma_t^{[kn]}$ in period $t > 1$ is chosen to minimize $d(u_t)^2 + \frac{1}{2}d(\frac{I_t}{Y_t})^2 + 0.01(\sigma_t^{[kn]})^2$, where $d(x_t) = (x_t^{\text{sim}} - \bar{x}_{t-1}^{\text{sim}}) - (x_t^{\text{data}} - \bar{x}_{t-1}^{\text{data}})$ is the difference between the deviation of the current simulated value x_t^{sim} from its prior geometrically weighted average $\bar{x}_{t-1}^{\text{sim}}$ and the corresponding empirical deviation. The weighted average is defined recursively as $\bar{x}_1^i = x_1^i$ and $\bar{x}_{t+1}^i = 0.95\bar{x}_t^i + 0.05x_t^i$ for $t > 1$, $i = \{\text{sim}, \text{data}\}$.

The first message of this simple experiment is that, when using TFP shocks as the source of disturbances to the economy, our model matches the standard business cycle moments of the data as well as a typical real business cycle model.[21]

We have seen before that our model is capable of creating asymmetric responses to shocks, as illustrated in figure 5. The force behind this effect are reallocation frictions,

that the cyclical variation of government expenditure and net exports combined approximately matches that of the investment share. This means that a simple model like ours, in which consumption and investment are the only components of output, can only match the variabilities of two out of the three variables output, consumption and investment.

[21]See for example Stadler (1994). Angeletos, Collard, and Dellas (2018) is an example of an attempt to replicate the co-movement of variables based on a different type of shock, higher-order belief dynamics in this case.

Table 2: Standard Deviations and Correlations

	data		model	
variable	standard deviation	correlation with output	standard deviation	correlation with output
real output	1.58	100	1.72	100
investment	7.21	86.6	7.25	98.3
consumption	1.21	76.9	0.43	74.2
employment	1.03	78.8	0.87	68.0
capital stock	0.36	32.5	0.42	-3.85
real interest rate	1.49	-5.61	0.24	92.9
unemployment rate	0.79	-87.2	0.81	-67.8
total factor productivity	1.11	89.4	1.40	94.9
labour productivity	0.99	77.0	1.29	86.9

All values in percent. Variables at quarterly frequency, 288 observations. Unemployment as rate, real interest as annualized rate, all other variables in logs. Standard deviations and correlations are for cyclical components, which were obtained with an HP_{1600} filter.
Data: Quarterly US data from 1948:Q1 to 2019:Q4. Output, consumption, investment: quarterly national account data (BEA). Employment and unemployment rate: monthly BLS data aggregated to quarterly frequency. Capital stock: annual volume estimates from BEA, interpolated to quarterly frequency using quarterly depreciation and investment figures from national accounts. Real interest rate: T-Bill rate adjusted for CPI inflation. TFP and labour productivity: calculated from output, employment and, in case of TFP, capital; assuming aggregate production technology $Y_t = A_t K_t^{0.33} N_t^{0.67}$ to calculate TFP A_t.
Simulation: Using the calibration summarized in table 1. Shock: Every period, a TFP shock drawn from a normal distribution with standard deviation 0.02 arrives with probability $\frac{1}{4}$. With probability $\frac{3}{4}$ there is no shock. The economy is simulated at quarterly frequency, starting in steady state. The first 96 periods are discarded.

search unemployment in this version of the model, which amplify investment-reducing shocks and boost investment increasing ones. Even with symmetric shocks driving fluctuations, we should therefore be able to see an asymmetry in the distribution of the cyclical components of output, unemployment and investment. Table 3 shows that these three variables are in fact skewed as predicted, both in the data and the model, with output and investment having a more pronounced tail on the negative side of the distribution and unemployment on the high end. For consumption, employment or interest rates, neither skewness measure is significantly different from zero in the data.

For all three of these variables and both coefficients reported in the table, the model produces the right signs and relative magnitudes of the skewness parameters. In fact, model output is slightly more asymmetric according to these measures with skewness values about

Table 3: The Chaban Moments of Skewness

variable	data Fisher (3rd moment)	data Pearson (median)	model Fisher (3rd moment)	model Pearson (median)
real output	-0.57	-0.13	-0.76	-0.38
	(0.27)	(0.24)	(0.37)	(0.14)
investment	-0.77	-0.32	-1.23	-0.43
	(0.28)	(0.23)	(0.42)	(0.12)
unemployment rate	0.70	0.38	0.96	0.78
	(0.34)	(0.16)	(0.35)	(0.21)

Fisher's moment coefficient of skewness and Pearson median skewness. Standard deviations (see footnote 18) in parentheses. See table 2 for details on data and simulation.

50 higher in absolute value than for US data. As all fluctuations are driven by a symmetric, in fact normal, shock, all the measured asymmetry is indeed endogenous. The majority of popular business cycle models, being locally linear in the shock, should be expected to have difficulties generating this kind of asymmetric response from symmetric shocks of moderate magnitude.

In our last exercise, we replace our rational, forward-looking households that maximize utility (6) under the constraint (9) with an ad-hoc specification of household behaviour. It is possible to close our model with a simple linear consumption function, as is done the Solow model, for example. We will use two specific functional forms for the households' consumption target. The first one makes planned consumption proportional to current output Y_t, just as in the Solow model. The second specification is an attempt to be more consistent with the permanent income hypothesis[22], making planned consumption proportional to trend output as a proxy for permanent income with a high propensity to consume, while assuming a much lower constant marginal propensity to consume out of deviations from this trend.[23] Whereas the first specification is purely static, the second one is history-dependent. Whenever it is not possible to satisfy the households' consumption plans due to reallocation constraints, we

[22] Friedman (1957).

[23] The exact model specifications are given in the notes to table 4.

continue to allow price changes to happen based on a static demand elasticity of θ^{-1} for the consumption good.

In this model version, shocks to the households' planned consumption result in a need for intersectoral reallocation, just like for example TFP shocks did in the case of forward-looking rational households. In fact, a shock that temporarily raises planned consumption and thus reduces investment triggers a recession followed by a recovery period, very similar to what would occur with forward-looking households. Table 4 reports the same correlation and skewness statistics for the two linear consumption functions under consideration – labelled "Solow" and "PIH" – as tables 2 and 3 did for the case of forward-looking households. These simulations are based on an additive shock to planned consumption $\sigma^{[c]}\bar{Y}$, where $\bar{Y}$ is trend output[24]. With a probability of 25% per quarter, $\sigma^{[c]}$ is drawn from a normal distribution with mean zero and standard deviation 2%, with 75% probability $\sigma^{[c]}$ is zero. The shock has a persistence of 50% per annum.

While the results are not as close to the data moments as those for the forward-looking households, they do capture the important patterns in the data and are generally not too far off.[25] This shows that the core mechanisms driving the co-movement of variables over the cycle are fairly robust with regards to the details of the specification of household behaviour. Moreover, this example demonstrates that it is possible to build a much simplified version of this business cycle model that does not involve forward-looking behaviour and thus substantially simplifies exposition, analysis and computation.

2.6 Alternative Types of Shocks

We have so far focused our attention on the effects of two types of shocks and in particular their ability to trigger recessions: a temporary TFP shock that impacts average productivity during its duration and a "marginal TFP shock," which only affects the productivity associ-

[24]See table 4 for a detailed specification.

[25]Remember that in this scenario, there are no productivity shocks. This explains why there is only minimal variation in measured TFP, and why labour productivity is relatively high during times of low output, when production is more capital intensive due to lower employment.

Table 4: Model Version with Ad-hoc Consumption Function

	model: "Solow" version		model: "PIH" version	
variable: 2$^{\text{nd}}$ moments	standard deviation	correlation with output	standard deviation	correlation with output
real output	0.70	100	0.67	100
investment	5.91	9.7	6.86	49.8
consumption	1.88	44.4	1.66	0.2
employment	1.01	98.5	0.96	97.5
capital stock	0.37	31.1	0.44	28.4
real interest rate	0.10	76.2	0.14	58.4
unemployment rate	0.93	-98.5	0.89	-97.5
total factor productivity	0.00	21.9	0.00	17.2
labour productivity	0.34	-86.0	0.34	-78.6
variable: skewness	Fisher (3$^{\text{rd}}$ moment)	Pearson (median)	Fisher (3$^{\text{rd}}$ moment)	Pearson (median)
real output	-1.30	-0.76	-0.89	-0.66
investment	-0.73	-0.09	-1.16	-0.25
unemployment rate	1.26	0.79	0.98	0.83

All standard deviations and correlations in percent. See tables 2 and 3 for details. *Simulation:* Using the calibration summarized in table 1. Planned consumption is given by $C_t = \kappa^{[avg]}\bar{Y}_t + \kappa^{[marg]}(Y_t - \bar{Y}_t) + C_t^{[\sigma]}$, $\bar{Y}_t = (1-\omega)\bar{Y}_{t-1}(1+\ell+g) + \omega Y_t$. We set $\omega = 5\%$, $\kappa^{[avg]} = \frac{7}{9}$, which results in the same nonstochastic steady state as in the model with forward-looking households, and $\kappa^{[marg]} = \frac{7}{9}$ in the "Solow" scenario and $\kappa^{[marg]} = 0.1$ in the "PIH" scenario. The shock component is updated according to $C_t^{[\sigma]} = \rho C_{t-1}^{[\sigma]} + \sigma_t^{[c]}\bar{Y}_t$, where the persistence ρ is 50% p.a. and the shock $\sigma_t^{[c]}$ arrives at a Poisson rate of one per year and is drawn from a mean-zero normal distribution with a standard deviation of 2%. The economy is simulated at quarterly frequency, starting in steady state. The first 96 periods are discarded.

ated with *additional* factor inputs. While the former concept has been a standard ingredient of macroeconomic models since Kydland's and Prescott's seminal work[26], we used the latter type of shock to see to what extent recession-like patterns can be caused by events that have no direct immediate impact on output or employment, but rather set in motion adjustment processes which then create the conditions for these variables to respond abruptly.

Ultimately, recessions in our model are caused by a sufficiently strong drop in demand for investment goods. Such a demand change can result from a variety of disturbances. This includes anything that changes the incentive of firms to invest or the desire of households

[26]Kydland and Prescott (1982)

to hold capital or financial wealth, even temporarily. It further includes events that make additional capital available in the market, such as increased project termination or firm exit, changes to the rate of depreciation or obsolescence or even trade disruptions.[27]

In what follows, we will briefly discuss how some of the shocks that are frequently relied upon in the macroeconomic literature, or at least were of interest at some point, map to our benchmark shock and to what extent they can trigger significant cyclical disturbances in our model.[28]

TFP Shocks TFP shocks are clearly the most widely used approach of including exogenous disturbances to macroeconomic models. Originating from the RBC framework, they have been used in most areas of macroeconomics, including monetary models and labour market theory. TFP shocks directly affect average productivity, D or A in our notation, and are typically assumed to be somewhat persistent.[29]

One way of understanding TFP shocks and the response of the economy to them is to realize that they change the productivity of the firm sector without changing factor input ratios accordingly, thus requiring adjustment. A negative TFP shock, for example, reduces productivity and with it the associated steady-state level of capital. However, the actual aggregate capital stock remains unchanged, making the economy more capital rich in relative terms. If the shock were to persist permanently, the full adjustment dynamics to lower the capital stock would be set in motion. In contrast to the *marginal* TFP shock we used before, however, TFP shocks do not merely trigger an adjustment. They come packaged with with an output response. In popular business cycle models, the magnitude and duration of output fluctuations is closely related to the amplitude and persistence of the TFP shock. As output

[27]The fact that, as we will see, a range of different shocks and disturbances can create similar fluctuations of output may be an opportunity to provide and test microfoundations for the TFP shock, which, while widely used, largely remains a black box.

[28]See Ramey (2016) for an overview of the main shocks relevant to the macroeconomic literature and their measurement. Smets and Wouters (2007) is an example of a paper including a number of different shocks in a single DSGE model.

[29]Kydland and Prescott (1982) use a highly persistent transitory shock instead of a permanent shock only for technical reasons, ibid., p. 1352, footnote 10.

fluctuations do not persist very long in the data, plausible TFP shocks must decay relatively quickly. Under rational expectations, this limits the extent and duration of the adjustment response. This close relationship between the magnitude and duration of the shock and the magnitude and duration of the endogenous response gives rise to the notion that RBC models and similar frameworks when combined with TFP shocks lack sufficient amplification and propagation mechanisms.[30]

How do TFP shocks play out in our model? Beyond their immediate effect in output, they do cause adjustments that require a reallocation between sectors, and can therefore initiate a recession event if this reallocation is strong enough. Due to their short persistence, an initial shift between sectors is soon reversed, causing further unemployment. The result is a recession-like event followed by a recovery period characterized by slowly declining unemployment. Because both increased unemployment and low TFP reduce output at the same time, negative TFP shocks can create large drops in output even with a moderate rise in unemployment.

It should be noted that logically, it might be possible to decouple the direct but short-lived productivity effect of a shock from a possibly more persistent incentive to adjust production patterns. Taking the oil price shocks as a plausible example of a shock that had a significant TFP component, it is not too far-fetched to consider the possibility that the abrupt changes in input prices triggered responses in the production sector that persisted even as the prices returned lower levels.

Monetary Policy Shocks Monetary policy is often considered to operate by setting nominal interest rates. In an environment in which frictions cause money to be nonneutral in the short run, monetary policy measures can have real effects for a limited time.

Both traditional Keynesian theory and a range of current monetary models highlight sticky prices as a source of monetary frictions. If the monetary authority changes the nominal interest rate, but prices cannot adjust quickly, the real interest rate has moved.

[30] See Rebelo (2005).

In our model, such a change of the real interest rate will affect saving and investment decisions. An excessively high real interest rate could cause a drop investment demand strong enough to create significant amounts of unemployment and trigger a recession.[31] Such an interest rate shock may have been a factor leading to the 1980 recession.

Fiscal Shocks Fiscal policy directly controls government expenditure, one of the components of GDP. In our model, abrupt changes to government expenditure can clearly trigger sharp rises in unemployment and recessions if they lead to inter-industry shifts in demand that cause increased project termination. It should be emphasized, though, that this immediate negative effect of government expenditure could arise both for increases and decreases in spending. The nature of the further adjustment following the shock would depend on what other components of GDP were crowded out or in.

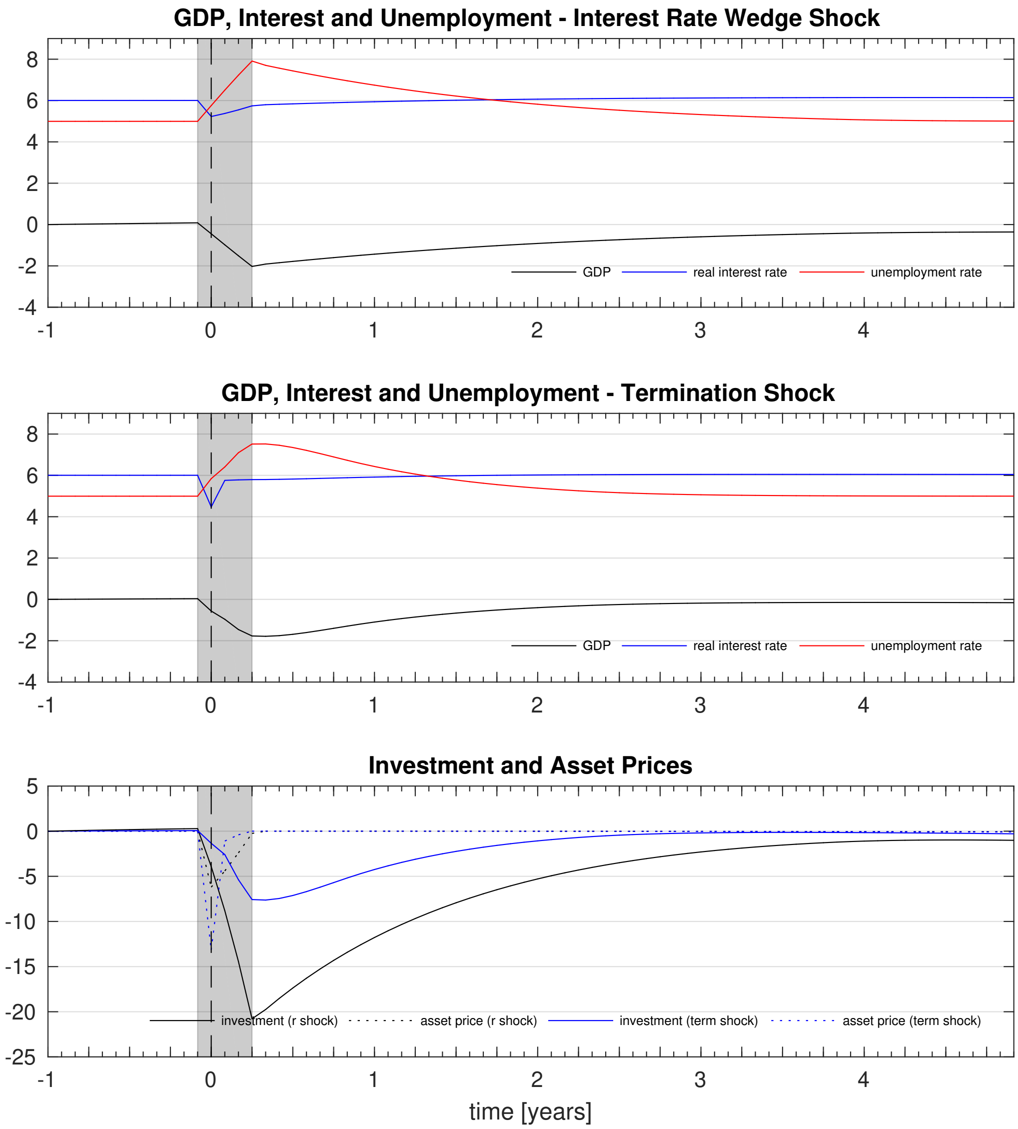
There is a range of other fiscal measures that have the potential to cause responses in output and employment. Changes in distortionary taxes or transfers that have a sufficiently large effect on household or firm behaviour are one example. Another one is the financing of government expenditure in regimes where the Ricardian equivalence fails to hold, so that government debt can crowd out private assets and affect the real interest rate.

Financial Shocks and Credit Crunches Following the 2008 recession, there has been increased interest in the role of financial shocks in causing disturbances to the macro economy.[32] As discussed before, in our model anything that reduces investment demand sufficiently to affect the project termination rate in the investment sector can cause a recession. Figure 10 shows two examples of such shocks that could arise due to frictions in financial markets.

[31] It is not possible to discuss the precise impact of monetary policy without being specific about the nature of money and monetary frictions. However, the interest rate wedge shock discussed below can serve as an example of how a disturbance to the interest rate relevant to the firm sector can trigger strong changes in unemployment and output.

[32] Examples include Jermann and Quadrini (2012), Benguria and Taylor (2020), and Görtz, Tsoukalas, and Zanetti (2022).

Figure 10: Full Model: Interest Rate Wedge and Project Termination Shocks



Note: Simulated impulse response. All ordinate values in percent. All rates annualized. Shock applied at time $t = 0$. All variables except interest rate and unemployment rate as logs detrended with BK_{96}^{96}.

Suppose that a situation arises in which, because of perceived credit risk, lenders become reluctant to provide funding to borrowers unless they receive a premium on the rate of return. This drives a wedge between the interest rate at which households save and the return firms are expected to generate on their assets. The top and bottom panels of figure 10 show the effect of such an interest wedge shock that initially creates a 6 percentage point wedge that shrinks at a rate of 50% per year. In the simulation, such a shock triggers a recession involving a large drop in investment.[33]

An alternative financial shock could be one in which a small number of firms become unable to finance their operations, as lenders are concerned about their creditworthiness due to them having certain unfavourable traits. Without being able to borrow, such firms cannot effectively compete in the market at their current scale and are forced to shut down projects. The bottom two panes of figure 10 show the effect of such a shock that initially raises the termination rate by 0.9 percentage points per month, but declines rapidly by 90% within the first year.[34] Again, such a shock triggers a recession, but one with a rather mild reduction in investment.

International Transmission of Cycles The possibility of international transmission of cycles has been an important issue in the post-war era. Earlier work often highlighted the role of trade flows[35], whereas more recently the focus has shifted towards effects linked to financial markets[36]. We have already identified financial shocks, which may be connected or transmitted internationally, as a possible source of cyclical fluctuations.

Similar to what has been said above about government expenditure, a large disruption

[33]The wedge only operates at the margin, affecting households' and firms' incentives to save and invest. All the interest paid by firms still contributes to the households' income, even if only part of it is tied to interest on savings.

[34]The cumulative effect is an additional project turnover of 4.6% in that year, compared to a base termination rate of 2.3% in steady state.

[35]The best-known comprehensive framework for studying transmission and policy in a traditional Keynesian open-economy setting was introduced by Mundell (1962) and Fleming (1962). Frankel and Rose (1998), Cravino and Levchenko (2016), and di Giovanni, Levchenko, and Mejean (2018) are examples of papers that highlight trade linkages for international patterns in business cycles.

[36]Recent examples include Ivashina, Scharfstein, and Stein (2015) and Miranda-Agrippino and Rey (2020).

to trade could in principle cause demand shifts across industries that are capable of starting a recession in open economies with a large volume of trade in our model. With increasing international integration of supply chains, it is easy to imagine how business cycles could become more synchronized due to trade effects.

Preference Shocks Shocks to the economy can in principle originate from the household sector, too. Anything that changes households' preferences for current versus future consumption and thus affects the target levels for the shares of consumption and investment in GDP will have similar effects as the interest rate wedge shock shown in figure 10.

Expectation-Driven Cycles It has been argued that the 2001 recession could be interpreted as the result of a correction in expectations following a boom built on unreasonably optimistic forecasts of future possible returns to investment that never materialized.[37]

There is a significant literature on the question of how and to what extent such episodes of expectation-driven cycles are possible in macroeconomic models built on the foundation of the RBC framework. The general problem is that in the absence of a contemporary productivity effect, to get the usual cyclical effects where investment and output move in the same direction, factor inputs have to increase during booms and fall during recessions. Yet, to the extent that labour supply is adjusted optimally to maximize the households' utility, income effects and consumption smoothing considerations tend to make households lower their labour supply on good news and increase it on bad news.[38]

Thanks to labour market frictions, cycles are associated with involuntary unemployment rather than efficient labour supply changes in our model, which means that they are less dependent on households' preferences. In fact, the marginal TFP shock we considered before

[37] See Karnizova (2012).

[38] The idea that expectations can be responsible for cycles dates back to at least Pigou (1927). Beaudry and Portier (2007) show that expectation-driven cycles cannot arise in one-sector real business cycle models. Jaimovich and Rebelo (2009) construct a rich model including variable utilization, adjustment costs and nonstandard preferences, in which expectation-driven cycles are possible. Karnizova (2010) assumes preferences over wealth. Also see Beaudry and Portier (2006) for an empirical investigation on how future TFP changes affect stock prices, and Christiano et al. (2008) for the relation to monetary policy.

has the flavour of an expectation shock, in the sense that it only affects the economy to the extent that project-level factor inputs are increased beyond their current level.

Figure 11 demonstrates the viability of expectation-driven cycles by showing the effect of a marginal TFP shock that is expected to only materialize with a delay of 5 years. The shock is exactly the same one as in figure 8. While the recession caused by the arrival of the news about the future shock is somewhat milder than if the effect were immediate, qualitatively the response is very much the same whether productivity effect applies now or later.

Figure 12 depicts a scenario that may be closer to to the late-1990s boom and early 2000s bust cycle. At time $t = 0$, expectations change such that a 4% marginal TFP shock is anticipated take effect three years later. At that time then, the positive effect does not materialize and agents realize that the anticipated productivity improvement will not happen.

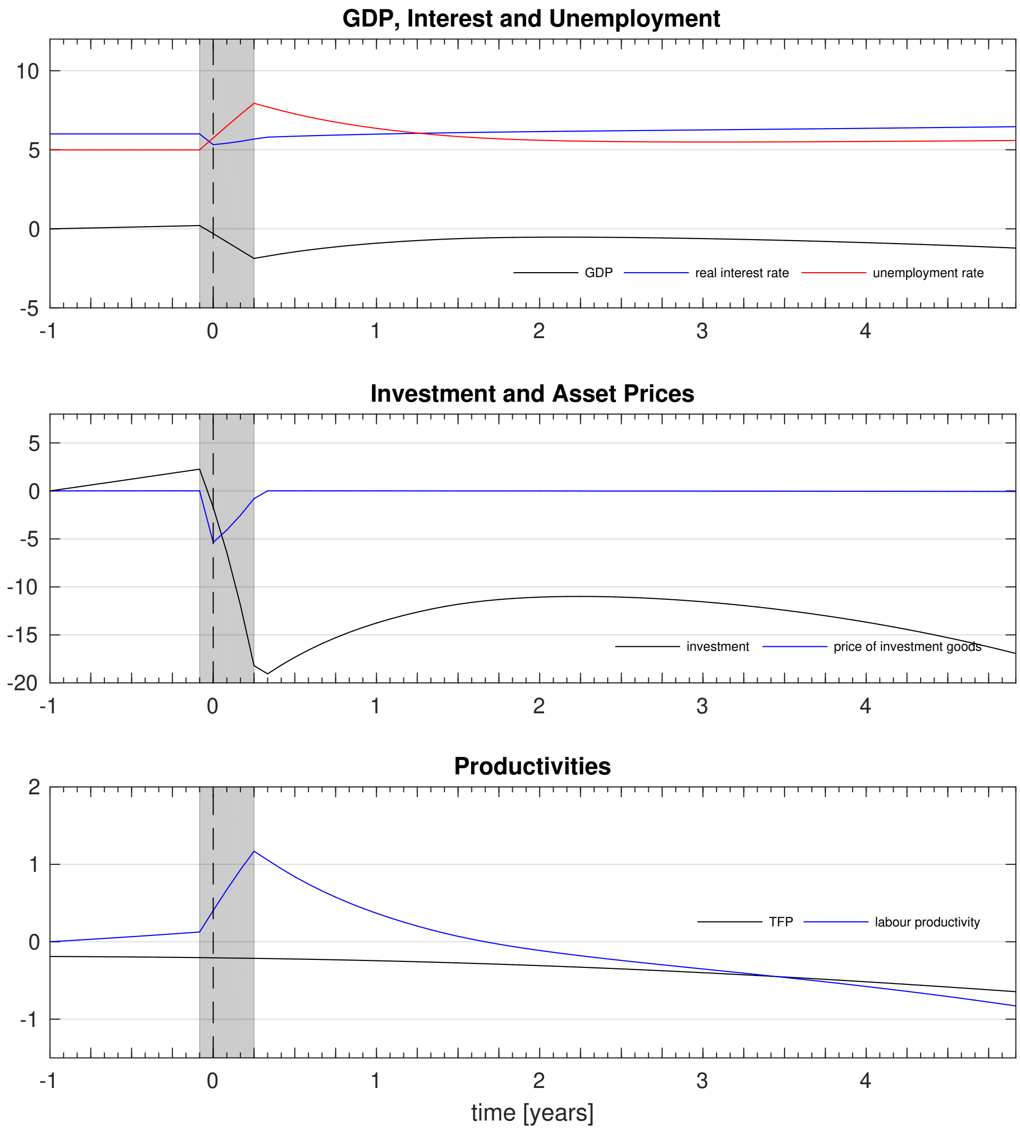
The immediate effect of the information shock is that resources are shifted towards the investment sector, resulting in a small increase in unemployment. Asset prices rise slightly as well. This is followed by an investment boom leading up to the expected productivity increase, in the course of which firm exit slows down sufficiently to make the unemployment rate drop below its steady state value, allowing output to rise above trend. Following the revelation that no productivity effect will occur, investment drops substantially, triggering a recession.

2.7 The Nature of Recessions

While being in line with the broad consensus that business cycles are the measurable response of the economic system to unexpected exogenous disturbances, the full model of section 2.5 provides some novel interpretations of what is driving recessions that are worth reiterating.

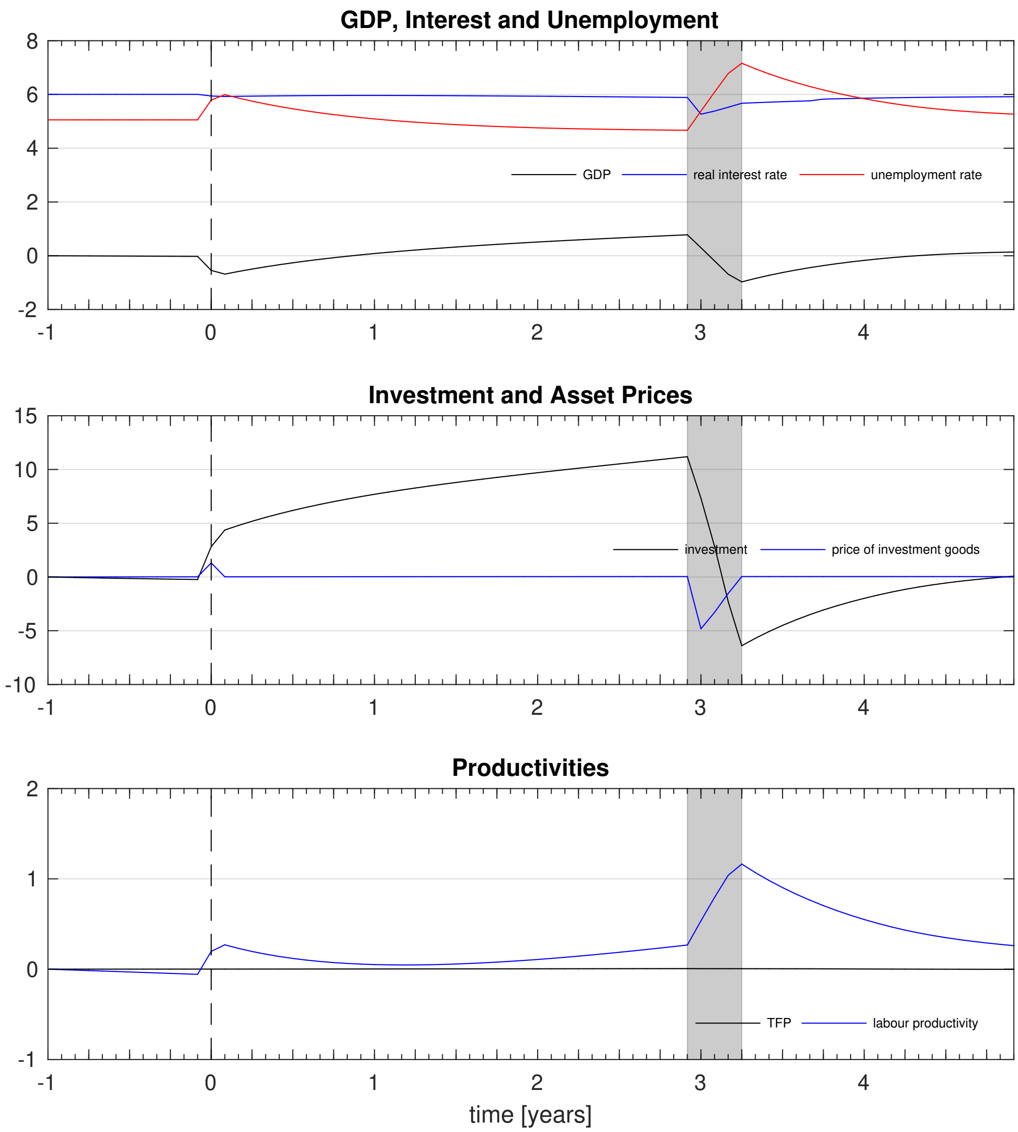
The *depth* of a recession is related to the changes in capital formation and investment caused by a shock. It is determined in part by an amplification effect, where a fundamentally necessary reduction in investment causes a fall in income and interest that reduces the

Figure 11: News about Shocks: Marginal TFP shock $\sigma^{[kn]} = 8\%$ that is announced at $t = 0$ and will take effect with a delay of 5 years



Note: Simulated impulse response. All ordinate values in percent. All rates annualized. News shock applied at time $t = 0$. All variables except interest rate and unemployment rate as logs detrended with BK_{96}^{96}.

Figure 12: Expectation-Driven Boom-Bust Cycle: Marginal TFP shock $\sigma^{[kn]} = 4\%$ that is expected to take effect after 3 years but does not materialize



Note: Simulated impulse response. All ordinate values in percent. All rates annualized. At time $t = 0$, expectations change such that a 4% improvement in marginal TFP is expected 3 years later. At the time when the productivity change is supposed to take effect, it unexpectedly does not materialize. All variables except interest rate and unemployment rate as logs detrended with BK_{96}^{96}.

demand for investment goods further. This amplification effect is also the cause of the *asymmetry* between recessions and investment booms.

The *duration* of recessions – the time it takes for the economy to move from peak to trough – as well as the drop in firm values compared to the price of consumption goods during the recession is the result of frictions that slow down the reallocation of factors between industries or sectors.

These mechanisms are fundamental to the model presented here, but the exact frictions through which they are achieved are not. It seems natural to include some form of labour market frictions in a business cycle model to allow for involuntary unemployment, and we leveraged the simple labour search model included for this purpose to both achieve the output reduction tied to intersectoral reallocation and limiting its rate at which it happens.

Additional or alternative frictions could change the effects quantitatively and make the model match certain aspects of the data better.[39] For example, it is not necessary to impose the same degree of search frictions on all types of labour. It seems plausible that these frictions mostly apply to certain types of specialized or skilled labour. Skilled workers may benefit more from a good match than unskilled workers, thus searching longer for new positions.[40] Firms that invest in the training of their labour force may be reluctant to lay off employees endowed with substantial amounts of valuable human capital. Both of these effects could contribute to the scarcity of skilled labour during reallocation processes, while complementary unskilled labour suffers involuntary unemployment due to low demand.

Set-up times for new projects or the capital used in these projects, whether newly produced or repurposed, would have very much the same effect in slowing down adjustments and temporarily lowering output as the labour market frictions we used. Moreover, informational or monetary frictions could reduce the rate at which relevant signals to adjust production

[39] The recessions generated by the model under labour market frictions only are on the shorter end of the spectrum of what is observed. Adding additional frictions that slow down factor reallocation further could help match the data better.

[40] Given that recessions are associated with intersectoral reallocation, mismatch unemployment as discussed in Şahin et al. (2014) is likely relevant for this particular group.

spread across industries, similarly delaying the response to shocks and thus prolonging and deepening recession events.[41]

2.8 The Relationship to the Neoclassical Growth Model

We have seen how it is possible to construct a business cycle model with attractive features simply by adding a very basic form of labour market frictions to the growth model introduced in part I of the paper.

The question we will ask now is: Would it not be possible to construct the same kind of business cycle model by directly building on the neoclassical growth model?

The answer to this question is: Absolutely. As we have argued before, when looked at from the angle of the aggregate economy, our model *is* the neoclassical growth model. What it adds is the underlying micro structure that explains and accounts for the productivity growth we observe at the aggregate level. This same structure formed the basis for the business cycle framework we explored. However, for the purpose of studying cycles, the fact that it determines long-term growth trends was not important. Instead, the way in which production is organized when looking under the hood of the aggregate growth model – in the form of small units that are each associated with a particular capital endowment, may be tied to a specific industry or sector, act competitively and can enter and exit markets freely – informed us how to think about the effects of shocks and the resulting cyclical patterns in aggregate data.

A translation of our business cycle model into the language of the neoclassical growth model would look like this. Aggregate output is determined by the constant returns to scale production function $F(K, N)$. Total factor productivity grows over time at a stable rate. Output is produced by atomistic firms under perfect competition. A unit mass of firms is associated with a capital stock k, which grows at the rate $\gamma^{[k]}$. The total mass of the firm

[41] The fact that in many past recessions, stock prices continued to decline for some time after an initial drop could be evidence that the magnitude or nature of the shock is not observed initially, or that relevant information about the required adjustment is slow to spread throughout affected parts of the economy.

sector is thus $\frac{K}{k}$. Each firm is specific to its sector. When firms exit for any reason, the labour associated with them becomes unemployed.

Aggregate output is given by $Y = C + I$, $C = \int_0^{m^{[C]}} F(k, n_i)di$ and $I = \int_{m^{[C]}}^{\frac{K}{k}} F(k, n_i)di$, where $m^{[C]}$ is the mass of firms in the consumption sector. All firms pay their factors according to their marginal products, and all factor markets clear, $N = \int_0^{\frac{K}{k}} n_i di$. If $\gamma^{[k]}$ is large enough, there will be an ongoing reduction in firm mass along a balanced growth path, which results in a corresponding flow of labour into unemployment.

A shock resulting in a reallocation of firms between sectors leads to additional unemployment, making production more capital intensive. This reduces the incentive to invest, thus amplifying shocks that already reduced production of investment goods, while dampening investment-increasing shocks. If direct hiring between firms is not possible, large enough shocks can lead to situations where a full shift of output between sectors is not immediately possible. The result is an outcome where the relative price of investment goods changes temporarily, while production techniques and wages diverge between groups of firms.

3 Broader Implications

By offering a different theoretical foundation of economic processes within the firm sector, our model can provide new insights into longstanding modelling issues, questions and puzzles in macroeconomics and finance. This section briefly discusses a few of those.

3.1 Demand-Determination of Output

A core tenet of Keynesian theories is that output and employment are, in some form, demand-determined. A traditional view would be that changes to autonomous components of any parts of aggregate demand – demand shocks in more contemporary language – have a direct quantity effect on output in the short term, as prices are fixed and supply accommodates demand. The secondary effect of such changes occurs in the labour market, where demand for workers reflects the scale of production. While the goods market is *assumed* to clear, the labour market is where the monetary frictions ultimately manifest themselves. The presumed stable relationship between output fluctuations and unemployment is given by Okun's law.

New Keynesian Theories in contrast highlight the role of price stickiness and expected price fluctuations for firms' optimal decisions. Monetary frictions can prompt the firm sector to respond to a variety of external shocks by adjusting output, leading to output gaps compared to a hypothetical frictionless economy.[42]

The cyclical patterns we observe in national account data are in line with the traditional Keynesian view. Variations in the major "endogenous" components of GDP, in particular investment and consumption, are positively correlated with output rather than showing a negative correlation with each other, supporting the view that price-related frictions prevent demand shocks in one sector to be offset by changes in another, thus resulting in a deviation of output from its potential level.

Our model is able to replicate these empirical patterns based on a mechanism that, at some level, is not unlike frictions highlighted by Keynesian theories. In both model frameworks, the inability of the system to immediately compensate for a demand change in in one part of the economy by changing output in other parts is at the core of aggregate business cycles. Of course, our mechanism directly relies on labour market frictions rather than on monetary phenomena such as price or wage rigidities to explain fluctuations in employment or GDP.

Yet, many of the empirical strategies used to identify demand shocks or structural features of the aggregate economy under inflexible prices would, at least to the extent that they pertain to real variables, pick up shocks and features of our model when applied to data generated from it. Consider for example applying the methodology introduced in the groundbreaking work of Shapiro and Watson (1988) and Blanchard and Quah (1989), which

[42]See Kiley (2013) for a discussion of output gap concepts.

classifies shocks based on their long-term effect, to time series generated from our model in the presence of permanent TFP shocks. The empirical method would identify these shocks correctly based on their permanent impact and classify them as supply shocks, and in addition to this interpret the strong transitory effects on output that result from large shocks due to amplification effects during recessions as temporary demand shocks.[43]

3.2 The Shape and Frequency of Recessions

Recessions look similar in that the rate at which unemployment rises between peak and trough does not differ much from cycle to cycle. This means that deeper recessions are typically also longer. In fact, for the eight recessions between 1960 and 2007, the rate of unemployment increases has ranged from 0.45 to 0.81 percentage points per quarter, and the correlation between the duration of these recessions and the overall change in the unemployment rate is 97%. Our model provides a simple explanation for this pattern. During a recession, supply in the relatively growing sector becomes very inelastic once a once a certain reallocation threshold is passed, as explained in section 2.5.2. This constrains the rate at which projects in the shrinking sector are shut down, and the the speed at which employment and output drop. A larger reallocation does not only result in a deeper and longer recession, but also a stronger temporary drop in firm values.

Categorizing recessions based on their shape is a popular part of the discussion of business cycles in the media and by policy makers. The distinction is often made between V-shaped and U-shaped recessions, where former are characterized by a swift recovery of output after the trough, whereas the latter refer to recession events where output stays depressed for longer. A glance at figure 1 makes it clear that there are distinctly different patterns, and that they are most easily identified based on unemployment trajectories, which are similar to but less noisy than output patterns. Following some recessions, unemployment drops rapidly, at least initially, almost at the same rate at which it had been rising during the

[43]See Ramey (2016) and Stock and Watson (2016) for an overview of VAR-based methodologies to identify shocks. Blanchard (1989) is an example of the estimation of a traditional Keynesian system.

recession, leading to a V-shape in output data. In other cases, the unemployment rate drops very gradually.

In our simulations, we focused on shocks that induce a lasting reallocation, which in turn causes the prolonged recoveries where elevated project termination rates lead to persistently high unemployment and thus U-shaped recessions. This suggests that in the case of V-shaped recessions, a shock triggers immediate intersectoral reallocation causing the recession, without resulting in significant further adjustments in the firm sector. One reason for this could be shocks that are strictly temporary and only result in a short-lived shift of demand between industries. Another one could be that an observed signal of a reallocation shock exaggerated the size of the required adjustment and thus led to a larger than necessary initial output and employment effect.

Another observation one might make is that recessions appear to be relatively evenly spaced,[44] which is atypical for events driven by a Poisson process. One possible explanation that our model suggests is that economies may become more vulnerable to shocks as the recovery from a recession progresses. Economies may simply be able to absorb shocks more easily while there is a lot of turnover anyway. In our model, the best indicator of this turbulence is the unemployment rate, and higher unemployment means larger labour market flows and thus a greater capacity to reallocate workers without additional job terminations should the need arise.

3.3 Unemployment

A large share of the variation in unemployment rates across countries and over time has been explained both theoretically and empirically based on institutional factors, micro and macroeconomic conditions, and their interaction. Our model can potentially contribute another puzzle piece by providing new microfoundations for the link between aspects of the macro state and unemployment outcomes.

[44] This is particularly true after the 1950s and when excluding the 1973 and 1980 recessions, which are tied to external political events and policy shocks.

By endogenizing project termination rates as one source of job turnover, the model can explain how unemployment is affected by adjustment processes that take place in the aggregate economy. In section 2.4, we focused on the persistently high unemployment during recoveries from a recession. Another example of a similar phenomenon are the astonishingly low unemployment rates in Japan and Germany between the 1950s and 1970s. In our model, it is possible to have unemployment rates persistently below the "natural rate" as long as firm termination rates are below their steady-state values. Both Germany and Japan saw an investment boom after the war, with investment as a share of GDP rising until it plateaued during the more turbulent 1970s. The observation that the end of these episodes of unusual investment growth coincides with a rise in the unemployment rates matches well with the explanation our model framework can provide.

3.4 Consumption Patterns

For plausible rates of risk aversion, theoretical models of consumption that formalize the notion of the permanent income hypothesis predict that consumption should be moderately responsive to changes in income. In the data, consumption is smoother than one would expect given the observed innovations to income, a phenomenon known as the excess smoothness puzzle.[45]

Our business cycle model of section 2.5 can contribute a few related points to the understanding of the co-movement of consumption and income. First, the volatility of consumption is limited by the ability of the firm sector to reallocate production between consumption and other types of output. Second, while output can change quickly in the short run due to labour market frictions, such changes may be tied to anticipated changes in long-term income that result from the shocks triggering the output fluctuations. Third, when judging consumption variation, it may be desirable to take the relative price of consumption q^{-1} into account.[46]

[45] See Deaton (1986) and Campbell and Deaton (1989).

[46] Taking the relative price of assets into account is likely – when done in isolation – to exacerbate the

In our model, shocks associated with recessions generate a *negative* co-movement between price-adjusted consumption and output, as it is the very reduction of desired saving in the household sector that triggers the recession with its steep decline in output. In equilibrium then, output ends up being fairly volatile while unadjusted consumption remains rather stable.

3.5 Asset Prices: Volatility

Asset pricing models predict that the value of assets is linked to future expected returns at any point in time. In the case of stock markets, it has been observed that the volatility of asset values is much higher than what could be justified based on actual fluctuations in dividends. This observation, first made by Shiller (1981) and LeRoy and Porter (1981) is known as the excess volatility puzzle.

Our approach offers one possible explanation for this excess volatility. Asset prices have two components, one related to the present value of the expected payoffs and one related to the relative price of capital goods relative to consumption goods. The excess volatility of asset prices would thus be attributed to these price changes.

In our model, dividends are given as the interest on capital net of internal financing to enable project-level growth. In the absence of firm or industry level risk, these dividends are extremely stable. Virtually all asset price and rate-of-return volatility is thus tied to changes in the price of capital q. While the simulation shown in figure 8 focuses on recessions and drops in equity prices, temporary increases in q during times of rapidly rising investment demand are equally possible.[47]

puzzle, by making adjusted consumption more countercyclical.

[47]Figure 12 shows an example of this in period $t = 0$.

3.6 Asset Prices: Equity Premium

The difficulty to rationalize the high average returns of equity compared to other forms of assets, the equity premium, has been one of the most high-profile problems in macroeconomics since the seminal paper of Mehra and Prescott (1985). At the core of the puzzle is the issue that given the riskiness of equity and consumption, the excess return on equity far exceeds the risk premia predicted by consumption-based asset pricing theories.

A variety of attempts have ben made to resolve this theoretical problem[48], but none of them has been generally accepted as a solution. It appears that there remains an unexplained component of about 3 to 4 percentage points of equity returns.[49]

While the models presented in this paper are not designed to address asset pricing issues, the characteristics of the capital price q introduced in section 2.5 provide some indications for where to look for explanations and where they are unlikely to be found. In the simulations reported in that section, we saw that a recession with a 3% drop in output below trend was associated with an abrupt but short-lived 10% drop in capital prices, while at the same time the real interest rate, measured as the return *per unit of assets*, did not change too dramatically.[50] If firms are financed through equity and (almost) risk-free debt, the drop in the value of equity will be more pronounced, while debt provides more stable returns. In the back-of-the-envelope calculations that follow, we will assume a debt-to-equity ratio of 1.5 (i.e. a ratio of equity to firm value of 0.4) and a frequency of recession events of once every eight years. This means that equity holders would expect a temporary drop in the values of their assets by one quarter to occur with a probability of $\frac{1}{8}$ per year. We will pretend that the variance of equity prices and rates of return outside of recessions is low enough in comparison to be ignored, and that the real risk-free interest rate is expected to remain

[48] Proposed solutions include rare events, alternative household preferences and incomplete markets, among others. See Mehra and Prescott (1988), Kocherlakota (1996), Siegel and Thaler (1997), Mehra (2003), and Mehra (2007) for assessments and reviews of the massive literature on this topic at different times.

[49] See Siegel (1992a), Siegel (1992b), and Dimson, Marsh, and Staunton (2008) for estimates.

[50] See figure 8 above.

approximately constant. With this setting in mind, I would like to make a few points about the equity premium.

The equity premium is unlikely to be a risk premium. This point may be provocative, given that a significant amount of research has been devoted to the question of how different preference structures can provide an explanation for the equity premium. My argument should be clear, however, once one takes into account the relative price of assets to consumption q instead of solely focusing on returns of assets. A stochastic version of the household's Euler equation (10) for an arbitrary asset i is given by

$$U'\left(\frac{C_t}{L_t}\right) = \mathbb{E}_t\left[\frac{1+r_t^{[i]}}{(1+\rho)(1+\ell)}\frac{q_{t+1}}{q_t}U'\left(\frac{C_{t+1}}{L_{t+1}}\right)\right].$$

Letting consumption C_{t+1} and population L_{t+1} be approximately nonstochastic, we arrive at the following relationship between the return on equity $i = e$ and the risk-free asset $i = f$ by eliminating the current marginal utility of consumption from the Euler equations associated with these assets.

$$\begin{aligned}
\frac{1+r_t^{[f]}}{(1+\rho)(1+\ell)}\mathbb{E}_t\left[\frac{q_{t+1}}{q_t}\right]U'\left(\frac{C_{t+1}}{L_{t+1}}\right) &= \frac{1}{(1+\rho)(1+\ell)}\mathbb{E}_t\left[(1+r_t^{[e]})\frac{q_{t+1}}{q_t}\right]U'\left(\frac{C_{t+1}}{L_{t+1}}\right) \Leftrightarrow \\
(1+r_t^{[f]})\mathbb{E}_t\left[\frac{q_{t+1}}{q_t}\right] &= \mathbb{E}_t\left[(1+r_t^{[e]})\frac{q_{t+1}}{q_t}\right] \Rightarrow \\
1 + r^{[f]} + \mathbb{E}_t\gamma^{[q]} &\doteq 1 + \mathbb{E}_t\left[r^{[e]} + \gamma^{[q]} + r^{[e]}\gamma^{[q]}\right] \Leftrightarrow \\
\mathbb{E}_t r^{[e]} &\doteq r^{[f]} - \mathbb{E}_t\left[r^{[e]}\gamma^{[q]}\right] \Rightarrow \\
\mathbb{E} r^{[e]} &\doteq \mathbb{E} r^{[f]} - \mathbb{E}\left[r^{[e]}\gamma^{[q]}\right]
\end{aligned}$$

The term $\mathbb{E}\left[r^{[e]}\gamma^{[q]}\right]$ is positive, as the appreciation of capital assets $\gamma^{[q]}$ is part of the rate of return $r^{[e]}$ on equity, so that these two variables are positively correlated. Based on these considerations, we would expect equity to carry a *negative* risk premium. Equity provides particularly high rates of return whenever capital values q are low but growing. The low value of q implies a high price of consumption. During recessions, the consumption level C of the representative household remains relatively stable, while the opportunity cost of consuming in terms of the household's wealth is very high. Equity can thus provide consumption insurance by providing high returns when there is a high demand for consumption goods.[51]

The equity premium could be a no-arbitrage condition. One important take-away message from the model of section 2.5 is that, while the crash of asset prices at the beginning of recessions is unpredictable, the following recovery is not. This view is supported by figure 1, which shows that stock prices recover following recessions, usually within a very short time.

Suppose that bonds and stocks had the same rate of return on average over the cycle. Savvy investors would want to hold short-term bonds most of the time, as they would protect them against the risk of capital price drops. Only during recessions, after stock prices have dropped substantially would they switch their portfolio to equity in order to take advantage of the high returns during the recovery of the stock market. This way they could earn excess returns by timing the market successfully.

This would, of course, be inconsistent with market clearing, as there would be virtually no demand for equity outside of recessions. In order for equity to be held in equilibrium, it must thus provide excess returns that compensate holders for the risk of stock market crashes. If, as per our example, there is a on-in-eight chance of stock prices dropping by one quarter in a year, the required premium is $\frac{1}{8}\frac{1}{4} = 3.125\%$, which is well in the ballpark of measured equity premia[52].

The old adage that it is impossible to time the market holds because of this excess return.

Another way of putting this is that when calculating the expected return of equity during times when the stock market performs normally, we should not use average historic returns that include predictable excess returns during recoveries. Once those are excluded, the

[51] Unrelated to the argument made here, Azeredo (2014) finds that given the characteristics of historic consumption growth data, the equity premium should be negative.

[52] Siegel (1992a), Siegel (1992b), and Dimson, Marsh, and Staunton (2008)

majority of the empirical equity premium may disappear.[53]

Why are real interest rates so stable? While it seems possible to rationalize the equity premium given the observed patterns in consumption and interest data, the overall patterns in asset markets are still hard to explain in the context of a representative-household model. The question is: If equity provides very high expected returns coming out of a recession, why is there still demand for bonds that yield significantly lower returns? The answer may lie in the heterogeneity of the household sector, where households with different characteristics acquire different assets for different purposes. One way of looking at this is that risk-free bonds are a claim to future consumption, whereas stocks are a claim to capital goods. While these are equivalent most of the time in the sense that they can be exchanged one for one, the equivalence temporarily breaks down during times of rapidly changing demand for consumption or investment goods that makes it impossible to attain market clearing without changes in the relative price. Still, the real puzzle seems to be: How can the risk-free interest rate be so low and stable in times when equity markets predictably do well.[54,55]

4 Conclusions

4.1 Summary

We have seen how changing the way in which we model the firm sector can help us both develop better models of endogenous technological progress and arrive a different understanding of the nature of cycles. Regardless of increasing returns from an atemporal perspective, a model economy can, at any point in time, exhibit constant returns to scale and function in many ways similar to the standard neoclassical growth model in the long

[53] Yet another way of looking at this is: the stochastic process driving rates of return matters, and our model suggests that it is not approximated well by a Markov process.

[54] This question appears to be related to the risk-free rate puzzle, Weil (1989).

[55] The returns on corporate bonds have been elevated during and following some of the recent recessions. While this could partially be an arbitrage effect related to high stock returns, it may equally plausibly be attributed to higher than usual perceived risk in the firm sector.

run.

In the short run, and in the presence of stochastic shocks, our model is capable of producing realistic business cycles. The familiar adjustment dynamics of the neoclassical growth model are modified by labour market frictions and firm-level processes. Any adjustment-related turbulence leads to additional unemployment, which amplifies investment-reducing dynamics while dampening investment increasing ones. Induced changes in project size and termination rates lead to above-average unemployment when investment is low, while also making possible investment booms that are accompanied by unusually low unemployment and above-trend output. Moreover, labour market frictions slow down reallocation processes following shocks, which can lead to recession episodes that go along with strong temporary changes in firm values.

Our model can explain why and how recessions happen, why they are relatively rare events, why there is no corresponding "positive" counterpart to recessions, why they are often followed by such a long, drawn-out recovery period and how the duration of this recovery is determined. We can understand the behaviour of a number of relevant aggregates and prices over the cycle, including GDP, investment, consumption, asset prices, the unemployment rate and the interest rate. These variables show business cycle patterns that are qualitatively in line with the data, and most of them match rather well quantitatively, even for the simple ad-hoc calibration used throughout this paper.

Overall, our model can help address most if not all of the issues listed in Sergio Rebelo's survey of the RBC and related literature[56] as "open questions in business cycle research:" The behaviour of asset prices, the cause of business cycles, alternatives to technology shocks and labour market patters.[57]

[56] Rebelo (2005)

[57] The remaining two challenges listed there are explaining the causes of the great depression and the industry-level co-movement of employment (Long and Plosser (1983)). Based on our model, the great depression as likely the deepest possible recession – investment dropped to almost zero in 1932/33 – could probably be explained as a combination of different shocks, including possibly the reversal of a prior investment boom based on overly optimistic assumptions.

4.2 Extensions and Modifications

The models presented in this paper are merely a first attempt at implementing some new ideas in logically consistent general equilibrium models. I have tried to keep the mechanics as simple as possible. There are many ways to improve these models, to make them more realistic or fit the data better. There are lots of alternative approaches that could be tried.

Exploring alternative or refined mechanisms that generate the frictions driving the slow reallocation of productive capacity between sectors and thus the asymmetry, depth and duration of recessions could be worthwhile. As discussed in section 2.7, allowing for different types of labour could help improve the aggregate characteristics of the model. Beyond that, it could contribute to the explanation of microeconomic phenomena, such as the different labour market experience of different qualities of workers or the hoarding of certain types of labour by firms. In the same way, frictions in the capital market could be investigated. Some capital may be industry or even firm specific, there could be costs associated with repurposing capital goods and it may take time to integrate additional capital into the production process. Apart from its effect on the cyclical properties of aggregates, all this could contribute to a better explanation of the patterns observed in capital markets.

In general, the framework of section 2 provides an interesting opportunity to integrate richer search models into a stochastic general equilibrium model. By endogenously determining the scale of the unit of production, our model has the right granularity to capture the interaction of households and firms in the labour market. Endogenizing the variation of matching and separation rates and unemployment durations with the state of the economy seems much more straightforward in our model than in a typical RBC or New Keynesian setting.

Informational or monetary frictions are another way of making the cyclical patterns generated by the model economy more realistic. The main reason for including monetary aspects is of course the study of monetary policy. In our setting, recessions are associated with significant welfare costs even in the absence of monetary frictions, making any policy measures that can dampen cycles all the more valuable.

Two aspects that are clearly missing from the model are the government and the international sector. Introducing fiscal policy will not only make it possible to assess the effect of aggregate government activity on the cyclicality of macroeconomic outcomes,[58]thanks to the endogeneity of project entry and exit that is at the core of business cycles, more subtile aspects of fiscal policy, such as the effect of distortionary taxes on project exit, can be studied as well.[59]

Most of these changes would make the model more complicated. It is also possible, however, to simplify the model. For illustrative purposes, the firm sector models introduced in both the growth section I-2 and the business cycle section 2 can be combined with a simple Solow-style saving equation as an ad-hoc stand-in for household behaviour, as seen in section 2.5.4. While this eliminates any dynamic choices in the resulting general equilibrium model – firm sector behaviour is statically determined already – it does retain all the interesting growth and business cycle dynamics.

4.3 Comparison with Alternative Models

While still relying on exogenous shocks, our model arguably explains cyclical fluctuations, and especially recessions, at a deeper level than real business cycle or new Keynesian models. In particular, it does not depend on the persistence of shocks to generate long-lasting output gaps, it features a strong internal amplification mechanism for adverse shocks and it can explain the asymmetry of cycles. Moreover, because a range of different shocks can result in the familiar business cycle patterns, it is not as vulnerable to many of the specific criticisms

[58]Depending on the nature of government expenditure, it may be possible to obtain interesting multiplier effects if shifts in government expenditure can attenuate the drop of investment at the onset of a recession.

[59]In addition to all this, a richer model of the household sector that allows for some degree of heterogeneity would be of interest. Beyond the immediate effects on consumption patterns and the interest rate, such an extension could help explain the demand for different types of assets and the differences in their returns. Household models with overlapping generations, borrowing constraints and uninsurable income risk are the natural candidates for this. Specifically, for the purpose of rationalizing the surprising return patterns of equity versus bonds, a setting with at least two types of agents with different risk exposure or access to financial markets may be the right direction to explore. Such models have been popular in the New Keynesian literature. See Galí (2018) for an overview of such "TANK" models.

of TFP shocks.

One important difference between our model and the RBC framework is that there are several sources of inefficiency in our setting. The most obvious one is the strong rise in unemployment during recessions. Another one that could be of interest is the duplication of R&D efforts, as a markets are supplied by a large number of projects when fewer could provide the same output at lower cost.

Since cycles are not primarily caused by contemporary changes of productivity in our model, but instead result from the reallocation of resources based purely on expectations about the future, the questions of if, how and under what circumstances cycles can be expectation-driven does not arise in the same way as it does in the RBC literature. Significant cyclical variation in macroeconomic variables exits as the result of recessions in our model, during which a drastic drop in output caused by frictions guarantees a plausible co-movement of output and investment.

Following the arguments put forward in the General Theory[60], traditional Keynesian models explain recessions as the result of the reduction in demand for output, combined with frictions, typically assumed to be monetary in nature, that lead the firm sector to adjust quantities rather than prices, preventing the labour market from clearing. In some ways, our model aligns with this description. In our setting, the reduction in demand for investment that is observed during recessions is caused by additional effective supply of existing assets. While Keynes argued that frictions elsewhere make a drop in demand for output or a component thereof turn into a reduction of output and hence employment, our story is that reduction in sectoral demand requires a reallocation, which, in the presence of labour market frictions, results in increased unemployment and thus lower output. While these two mechanisms are different, it is interesting to note that – given that we do not directly observe concentration, project termination and the reallocation of productive assets, which are at the core of recessions in our model, in the national accounts – both theories are

[60]Keynes (1936).

observationally rather similar.

Finally, it is worth noting that the amplification effect in our model that leads to recessions being triggered in the case of large enough shocks is somewhat reminiscent of a negative multiplier-accelerator effect. A drop in investment demand results in a deterioration of aggregate conditions, which in turn prompts an even stronger reduction in investment.

4.4 Concluding General Remarks

In this paper, I have presented a general equilibrium model with some novel features that arguably has the potential to fit certain patterns in macroeconomic data better than existing theories. This includes its ability to produce stable growth rates in a robust way and in changing environments as well as to show cyclical behaviour and in particular realistic recessions. In doing so, it provides us with new explanations of how economies grow and what causes economic downturns and spikes in unemployment.

Like most contemporary economic models, it makes very specific assumptions, many of which could be questioned and criticized for good reasons. I would therefore like to take a step back from the concrete models presented here and ask: What are the important core elements of our story, what are the general lessons that can be learned?

The mechanism at the heart of all the new and interesting features of our model is a competitive market organization which ensures that firms operating at increasing returns to scale with regards to their long-term production potential adjust project sizes and market shares in equilibrium in such a way that at any point in time at the level of an industry or national economy, aggregate production patterns are consistent with constant returns to scale. For this to be possible, entry or exit always needs to be an option for firms.

Constant returns to scale are the standard model assumption in macroeconomics for many good reasons. They align with our experience with national economies and aggregate data, they make models scale-invariant and generally well-behaved, and they simply make sense logically when considering production at a large scale. Yet, there are good reasons to deviate from this assumption when modelling individual firms. As has been emphasized in the literature on endogenous growth, if ideas, technologies or blueprints have value and affect productivity regardless of the quantity of output produced, they contribute to increasing returns to scale. Especially for industries characterized by significant expenditures on R&D, the assumption of constant average costs at the firm level does therefore not appear particularly compelling.

It is important to remember that though being a convenient, widely used and often helpful assumption, the existence of a largely time-invariant aggregate production function with constant returns to scale and fixed labour shares is hardly a law of nature. If it is a consistent and stable feature of national economies, we should consider it an emergent property that is worth explaining. The Lucas critique reminds us that in order to predict how an economic systems responds to previously unseen events, we need to employ a model structure that goes deeper than what is needed to superficially replicate the familiar. The same idea applies to the macroeconomic production function. In order to understand the mechanisms that drive *dynamic* outcomes such as ongoing productivity growth or recessions, we cannot rely on an ad-hoc structure that was chosen because it is consistent with *static* features of what we observe.

It is therefore time to move beyond the largely static notion of aggregate production possibilities and crack open the black box that is the macroeconomic production function to expose the microeconomic mechanisms that are inside it. In doing so, we may learn things we did not expect to learn and solve puzzles we did not expect to solve.

Macroeconomic models, especially those dealing with business cycle phenomena and the corresponding macroeconomic policies, have become increasingly complex. However, when investigating patterns that are extremely robust and stable in the sense that they arise in similar forms in a variety of circumstances and repeatedly over long periods of time, one would expect or at least hope that a conceptually simple explanation exists that is similarly robust and stable as the observed phenomenon. The simple recurring patterns of recessions and the incredibly stable average rate of productivity growth are two such phenomena, and it should be possible to adequately explain and model them at least qualitatively without having to rely on exceedingly complex model structures and very particular calibration choices.

I hope this paper has made the point that despite the broad and deep understanding we have of many macro forces shaping our economic experience, looking for simpler and possibly deeper explanations of what we observe can still be a fruitful endeavour.

References

Abraham, Katharine G. and Lawrence F. Katz (1986). "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?" In: *Journal of Political Economy* 94.3, pp. 507–522. DOI: 10.2307/1833046.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas (2018). "Quantifying Confidence". In: *Econometrica* 86.5, pp. 1689–1726. DOI: 10.3982/ECTA13079. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA13079.

Azeredo, Francisco (2014). "The equity premium: a deeper puzzle". In: *Annals of Finance* 10.3, pp. 347–373. DOI: 10.1007/s10436-014-0248-7.

Baxter, Marianne and Robert G. King (1999). "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series". In: *The Review of Economics and Statistics* 81.4, pp. 575–593. DOI: 10.2307/2646708.

Beaudry, Paul and Franck Portier (Sept. 2006). "Stock Prices, News, and Economic Fluctuations". In: *American Economic Review* 96.4, pp. 1293–1307. DOI: 10.1257/aer.96.4.1293.

— (July 2007). "When can changes in expectations cause business cycle fluctuations in neo-classical settings?" In: *Journal of Economic Theory* 135.1, pp. 458–477. DOI: 10.1016/j.jet.2006.06.009.

Benguria, Felipe and Alan M. Taylor (Dec. 2020). “After the Panic: Are Financial Crises Demand or Supply Shocks? Evidence from International Trade”. In: *American Economic Review: Insights* 2.4, pp. 509–26. DOI: 10.1257/aeri.20190533.

Blanchard, Olivier Jean (1989). “A Traditional Interpretation of Macroeconomic Fluctuations”. In: *The American Economic Review* 79.5, pp. 1146–1164. DOI: 10.2307/1831442.

Blanchard, Olivier Jean and Danny Quah (1989). “The Dynamic Effects of Aggregate Demand and Supply Disturbances”. In: *The American Economic Review* 79.4, pp. 655–673. DOI: 10.2307/1827924.

Caballero, Ricardo and Mohamad L Hammour (1994). “The Cleansing Effect of Recessions”. In: *American Economic Review* 84.5, pp. 1350–68.

Campbell, John and Angus Deaton (July 1989). “Why is Consumption So Smooth?” In: *The Review of Economic Studies* 56.3, pp. 357–373. DOI: 10.2307/2297552. eprint: https://academic.oup.com/restud/article-pdf/56/3/357/4611159/56-3-357.pdf.

Christiano, Lawrence et al. (Oct. 2008). *Monetary policy and stock market boom-bust cycles.* Working Paper Series 955. European Central Bank.

Cooper, Russell and John Haltiwanger (1993). “The Aggregate Implications of Machine Replacement: Theory and Evidence”. In: *The American Economic Review* 83.3, pp. 360–382. DOI: 10.2307/2117523.

Cooper, Russell, John Haltiwanger, and Laura Power (Sept. 1999). “Machine Replacement and the Business Cycle: Lumps and Bumps”. In: *American Economic Review* 89.4, pp. 921–946. DOI: 10.1257/aer.89.4.921.

Cravino, Javier and Andrei A. Levchenko (Nov. 2016). “Multinational Firms and International Business Cycle Transmission”. In: *The Quarterly Journal of Economics* 132.2, pp. 921–962. DOI: 10.1093/qje/qjw043. eprint: https://academic.oup.com/qje/article-pdf/132/2/921/30637482/qjw043.pdf.

Deaton, Angus (Apr. 1986). *Life-Cycle Models of Consumption: Is the Evidence Consistent with the Theory?* NBER Working Papers 1910. National Bureau of Economic Research, Inc.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Mejean (2018). "The Micro Origins of International Business-Cycle Comovement". In: *The American Economic Review* 108.1, pp. 82–108. DOI: 10.2307/26527946.

Dimson, Elroy, Paul Marsh, and Mike Staunton (2008). "The Worldwide Equity Premium: A Smaller Puzzle". In: *Handbook of the Equity Risk Premium*. Elsevier. Chap. 11, pp. 467–514. DOI: 10.1016/B978-0-444-50899-7.X5001-5.

Fleming, M. J. (1962). *Domestic Financial Policies Under Fixed and Floating Exchange Rates*. IMF Staff Papers 9(3), pp. 369–379.

Frankel, Jeffrey A. and Andrew K. Rose (1998). "The Endogeneity of the Optimum Currency Area Criteria". In: *The Economic Journal* 108.449, pp. 1009–1025. DOI: 10.2307/2565665.

Friedman, Milton (1957). *A Theory of the Consumption Function*. Princeton: Princeton University Press.

Galí, Jordi (Aug. 2018). "The State of New Keynesian Economics: A Partial Assessment". In: *Journal of Economic Perspectives* 32.3, pp. 87–112. DOI: 10.1257/jep.32.3.87.

Görtz, Christoph, John D. Tsoukalas, and Francesco Zanetti (Oct. 2022). "News Shocks under Financial Frictions". In: *American Economic Journal: Macroeconomics* 14.4, pp. 210–43. DOI: 10.1257/mac.20170066.

Hodrick, Robert J. and Edward C. Prescott (1997). "Postwar U.S. Business Cycles: An Empirical Investigation". In: *Journal of Money, Credit and Banking* 29.1, pp. 1–16. DOI: 10.2307/2953682.

Hosios, Arthur J. (1994). "Unemployment and Vacancies with Sectoral Shifts". In: *The American Economic Review* 84.1, pp. 124–144. DOI: 10.2307/2117974.

Ivashina, Victoria, David S. Scharfstein, and Jeremy C. Stein (2015). "Dollar Funding and the Lending Behavior of Global Banks". In: *The Quarterly Journal of Economics* 130.3, pp. 1241–1282. DOI: 10.2307/26372626.

Jaimovich, Nir and Sergio Rebelo (Sept. 2009). "Can News about the Future Drive the Business Cycle?" In: *American Economic Review* 99.4, pp. 1097–1118. DOI: 10.1257/aer.99.4.1097.

Jermann, Urban and Vincenzo Quadrini (Feb. 2012). "Macroeconomic Effects of Financial Shocks". In: *American Economic Review* 102.1, pp. 238–71. DOI: 10.1257/aer.102.1.238.

Karnizova, Lilia (2010). "The spirit of capitalism and expectation-driven business cycles". In: *Journal of Monetary Economics* 57.6, pp. 739–752. DOI: 10.1016/j.jmoneco.2010.06.002.

— (2012). "News Shocks, Productivity and the U.S. Investment Boom-Bust Cycle". In: *The B.E. Journal of Macroeconomics* 12.1. DOI: 10.1515/1935-1690.2436.

Keynes, John M. (1936). *The General Theory of Employment, Interest and Money*. Macmillan and Co.

Kiley, Michael T. (2013). "Output gaps". In: *Journal of Macroeconomics* 37, pp. 1–18. DOI: 10.1016/j.jmacro.2013.04.002.

Kocherlakota, Narayana R. (Mar. 1996). "The Equity Premium: It's Still a Puzzle". In: *Journal of Economic Literature* 34.1, pp. 42–71.

Kydland, Finn E. and Edward C. Prescott (1982). "Time to Build and Aggregate Fluctuations". In: *Econometrica* 50.6, pp. 1345–1370. DOI: 10.2307/1913386.

LeRoy, Stephen F. and Richard D. Porter (1981). "The Present-Value Relation: Tests Based on Implied Variance Bounds". In: *Econometrica* 49.3, pp. 555–574. DOI: 10.2307/1911512.

Lilien, David M. (1982). "Sectoral Shifts and Cyclical Unemployment". In: *Journal of Political Economy* 90.4, pp. 777–793. DOI: 10.2307/1831352.

Long John B, Jr and Charles I Plosser (Feb. 1983). “Real Business Cycles”. In: *Journal of Political Economy* 91.1, pp. 39–69. DOI: 10.1086/261128.

Mehra, Rajnish (2003). *The Equity Premium: Why is it a Puzzle?* NBER Working Papers 9512. National Bureau of Economic Research, Inc.

— (2007). “The Equity Premium Puzzle: A Review”. In: *Foundations and Trends in Finance* 2.1, pp. 1–81. DOI: 10.1561/0500000006.

Mehra, Rajnish and Edward C. Prescott (Mar. 1985). “The equity premium: A puzzle”. In: *Journal of Monetary Economics* 15.2, pp. 145–161.

— (1988). “The equity risk premium: A solution?” In: *Journal of Monetary Economics* 22.1, pp. 133–136. DOI: 10.1016/0304-3932(88)90173-0.

Miranda-Agrippino, Silvia and Hélène Rey (May 2020). “U.S. Monetary Policy and the Global Financial Cycle”. In: *The Review of Economic Studies* 87.6, pp. 2754–2776. DOI: 10.1093/restud/rdaa019. eprint: https://academic.oup.com/restud/article-pdf/87/6/2754/36143045/rdaa019.pdf.

Mundell, Robert A. (1962). *The Appropriate Use of Monetary and Fiscal Policy for Internal and External Stability*. IMF Staff Papers 9(1), pp. 70–79.

Pigou, Arthur Cecil (1927). *Industrial Fluctuations*. London: MacMillan and Co.

Ramey, Valerie (2016). “Macroeconomic Shocks and Their Propagation”. In: vol. 2. Elsevier. Chap. Chapter 2, pp. 71–162.

Rebelo, Sergio (June 2005). “Real Business Cycle Models: Past, Present and Future”. In: *Scandinavian Journal of Economics* 107.2, pp. 217–238. DOI: 10.1111/j.1467-9442.2005..

Şahin, Ayşegül et al. (Nov. 2014). “Mismatch Unemployment”. In: *American Economic Review* 104.11, pp. 3529–64. DOI: 10.1257/aer.104.11.3529.

Shapiro, Matthew D. and Mark W. Watson (1988). “The Sources of Business Cycle Fluctuations”. In: *NBER Macroeconomics Annual 1988, Volume 3*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 111–156.

Shiller, Robert J. (1981). "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" In: *The American Economic Review* 71.3, pp. 421–436. DOI: 10.2307/1802789.

Siegel, Jeremy J. (1992a). "The Equity Premium: Stock and Bond Returns since 1802". In: *Financial Analysts Journal* 48.1, pp. 28–46. DOI: 10.2307/4479502.

— (Apr. 1992b). "The real rate of interest from 1800-1990 : A study of the U.S. and the U.K". In: *Journal of Monetary Economics* 29.2, pp. 227–252.

Siegel, Jeremy J. and Richard H. Thaler (Mar. 1997). "Anomalies: The Equity Premium Puzzle". In: *Journal of Economic Perspectives* 11.1, pp. 191–200. DOI: 10.1257/jep.11.1.191.

Smets, Frank and Rafael Wouters (June 2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". In: *American Economic Review* 97.3, pp. 586–606. DOI: 10.1257/aer.97.3.586.

Stadler, George W. (1994). "Real Business Cycles". In: *Journal of Economic Literature* 32.4, pp. 1750–1783. DOI: 10.2307/2728793.

Stock, James H. and Mark W. Watson (1999). "Chapter 1: Business Cycle Fluctuations in US Macroeconomic Time Series". In: vol. 1. Handbook of Macroeconomics. Elsevier, pp. 3–64. DOI: 10.1016/S1574-0048(99)01004-6.

— (Dec. 2016). "Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics". In: *Handbook of Macroeconomics*. Ed. by J. B. Taylor and Harald Uhlig. Vol. 2. Handbook of Macroeconomics. Elsevier. Chap. 0, pp. 415–525. DOI: 10.1016/bs.hesmac.2016.04.

Weil, Philippe (1989). "The equity premium puzzle and the risk-free rate puzzle". In: *Journal of Monetary Economics* 24.3, pp. 401–421. DOI: 10.1016/0304-3932(89)90028-7.