Abstract
We investigate UK labour productivity over the long run, comparing it to other advanced economies and focus on the sharp slowdown since the global financial crisis. Using a growth accounting framework, we find a dominant role for total factor productivity (TFP), but it turns out that capital shallowing is also important. Two macroeconomic trends deepen the puzzle of this slowdown. There has been a decline in real interest rates over the past 30 years and an increase in labour supply since 2008, both of which ought to have increased investment. And yet, the ratio of (nominal) private and public investment to GDP has fallen over time. We go on to examine the UK’s productivity performance through the lens of standard neoclassical models and reconsider the secular stagnation debate in the UK context. Finally, we survey several explanations from recent economic literature for the poor performance of investment.

Reference Details
CWPE 2437
Published 27 June 2024

Key Words Productivity, Macroeconomics, Real Interest Rates, Investment
JEL Codes E22, E24, E32, E44, E51, E62, O16, O42, O47

Website www.econ.cam.ac.uk/cwpe
Abstract:

We investigate UK labour productivity over the long run, comparing it to other advanced economies and focus on the sharp slowdown since the global financial crisis. Using a growth accounting framework, we find a dominant role for total factor productivity (TFP), but it turns out that capital shallowing is also important. Two macroeconomic trends deepen the puzzle of this slowdown. There has been a decline in real interest rates over the past 30 years and an increase in labour supply since 2008, both of which ought to have increased investment. And yet, the ratio of (nominal) private and public investment to GDP has fallen over time. We go on to examine the UK’s productivity performance through the lens of standard neoclassical models and reconsider the secular stagnation debate in the UK context. Finally, we survey several explanations from recent economic literature for the poor performance of investment.

Keywords: Productivity, Macroeconomics, Real Interest Rates, Investment.

1 Introduction

The material improvement in living standards, as measured by the increase in the production of goods and services, has been an artefact of the modern world. Britain has been characterised as the first industrial nation (Mathias, 1969), and accordingly, with a sustained period of economic growth, it reached a peak in 1900 of 9.4% of world output (Maddison, 2010). This fraction has declined with the increasing growth of emerging parts of the world but also because Britain’s relative performance has tended to deteriorate. At present, the UK is at some 2.3% of world GDP and seems likely to fall to 2% or less over the next 20 years. That observation does not necessarily imply that living standards, as measured broadly by per capita income, will fall but more that many other nations may see their living standards rise materially faster. This process has been emphasised by the phenomenon that has come to be known as the productivity puzzle: growth in measured productivity has fallen far behind previous trends and this has opened a large gap between anticipated and actual income per head (see Figure 1).

Figure 1: Recent trends in GDP per head in the UK.

![Graph showing recent trends in GDP per head in the UK.]

Note: Real GDP per capita in the UK expressed in 2019 thousands of British pounds. Source: NIESR; NIESR (2022).

1Data from NiGEM database, accessed October 2022, NIESR’s July 2022 forecast and NiGEM stochastic simulations; NIESR (2022).
We might characterise "normal times" as involving the steady expansion of the economy's supply capacity with small jolts or shocks to demand, from changes in confidence, sentiment or from overseas markets, that lead to small fluctuations around that expansion path: these small fluctuations are called business cycles. The growth in capacity can most easily be thought about as the sum of the growth in inputs, typically capital and labour, and how they combine to produce a given level of output, total factor productivity: a residual representing technological progress, changes in management practices, organizational change, network effects and other factor affecting production that are not captured by our measures of labour and capital. The growth in total factor productivity is thus simply the difference between the growth rate in the index of outputs and the growth rate in the index of inputs in production and does not per se explain productivity performance. Providing short-run fluctuations in demand, which take the economy temporarily away from this level of supply capacity, do ebb away then we can concentrate on accounting for economic growth in terms of these factors alone. But such accountancy exercises, while useful and widely deployed, do not provide any causal measure and thus are not a guide to the appropriate policy interventions to reverse observed trends.

The problem following the financial crisis of 2007-8 is not so much that we cannot account for growth in output in this manner, but more that the growth in output seems concentrated in the increase of inputs, especially labour, rather than productivity. We document a marked increase in aggregate hours worked following the financial crisis, mostly driven by increased participation of older workers. And so, labour productivity, as measured by output per hour worked, stalled after 2008. It is as though the economy rather than working harder and smarter has rather been spending more time at the office.

This paper investigates some important questions regarding productivity in the United Kingdom from a macroeconomic perspective. We start by examining the UK’s recent and long-term productivity trends and contrast these against trends in comparable economies (section 2). We then apply modern growth accounting techniques to examine the recent productivity slowdown in the UK and show that the results and their interpretation are impacted by the choice of methodology. Section 3 examines recent UK productivity performance under the lens of standard neoclassical growth models. The 'secular stagnation' debate is reviewed in section 4. This section also focuses on the UK’s labour demographics and comments on the country’s investment dynamics to illustrate some of the arguments emanating from the 'secular stagnation' literature. The UK economy suffers from chronic under-investment despite real interest rates trending down. We examine this puzzle in section 5 and comment on some of the explanations brought by recent economic literature. Section 6 argues that aggregate total factor productivity is not necessarily an exogenous process that remains out of the remit of economic policy and that there is scope to improve
productivity growth through long-term policy commitments. The issue of government debt, fiscal policy and their implications for productivity growth are discussed in section 7, and section 8 concludes.

2 Productivity trends in the United Kingdom

2.1 Long-term trends of the UK labour productivity and the recent slowdown

By labour productivity, most economists refer to output per hour worked. We will adopt this convention in the remainder of this work. We use the Bank of England millennium spreadsheet (Thomas and Dimsdale, 2017) from 1760 to 2016 and complement it using ONS data from 2017 to 2021 to construct an output per hour series for the UK covering the 1760-2021 period. The resulting time series in Figure 2 shows the exceptional period of labour productivity growth after the second world war. In this so-called 'golden age' period (1946 to 1974), productivity grew by an average of 3.7% a year (Table 1). Labour productivity growth slowed down after the oil shock (1973) but remained robust at around an average of 2.3% a year until the onset of the global financial crisis (GFC). Very low levels of productivity growth have plagued the post-GFC period (2008-2021). This latest episode of productivity slowdown stands out as unprecedented since the onset of the industrial revolution (Crafts, 2021).

Table 1: Long-term trends in the growth of labour productivity, UK.

<table>
<thead>
<tr>
<th>Period</th>
<th>Growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1760-1790</td>
<td>-0.40%</td>
</tr>
<tr>
<td>1790-1840</td>
<td>0.60%</td>
</tr>
<tr>
<td>1840-1873</td>
<td>1.50%</td>
</tr>
<tr>
<td>1873-1914</td>
<td>0.90%</td>
</tr>
<tr>
<td>1914-1945</td>
<td>1.40%</td>
</tr>
<tr>
<td>1945-1974</td>
<td>3.70%</td>
</tr>
<tr>
<td>1974-2008</td>
<td>2.30%</td>
</tr>
<tr>
<td>2008-2020</td>
<td>0.50%</td>
</tr>
</tbody>
</table>

Note: Average annual growth rates of the UK GDP/hours Index (2013=100). Data up to 2016 from Thomas and Dimsdale (2017), ONS data has been used for the remaining years (July 2022 data).

Another possible definition is output per worker. However, this definition ignores intensive labour supply, thus showing an improvement in productivity when workers increase labour input by supplying more hours.
Revisions to national account data affect productivity measures, thus affecting the measured magnitude of the post-GFC slowdown (Table 2). Using data from the Blue Book 2013, the average annual growth of output per hour worked dropped from 2.4% for the 1997-2007 period to 0.1% for the 2009-2013 period, implying a slowdown of 2.3% in the average labour productivity growth rate. When using 2021 blue book data, the average growth rate of output per hour worked between 1997 and 2007 is revised down to 2%, thus implying that the average growth rate of labour productivity has slowed only by 1.3% after 2008. Nevertheless, the post-GFC slowdown remains substantial for all available data vintages and seems unlikely to be revised away.

2.2 International comparisons

Figure 3 compares the long-term labour productivity trends in the United Kingdom, the United States, France and Germany between 1890 and 2019. The figure shows that the UK economy remained more productive than the US economy until the First World War. The
Table 2: Average annual growth rates of output per hour worked, UK.

<table>
<thead>
<tr>
<th>Blue Book</th>
<th>1997 to 2007</th>
<th>2009 to latest year</th>
<th>Change in growth rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>2.1%</td>
<td>1.1%</td>
<td>-1.0%</td>
</tr>
<tr>
<td>2012</td>
<td>2.5%</td>
<td>1.3%</td>
<td>-1.2%</td>
</tr>
<tr>
<td>2013</td>
<td>2.4%</td>
<td>0.1%</td>
<td>-2.3%</td>
</tr>
<tr>
<td>2014</td>
<td>2.2%</td>
<td>0.5%</td>
<td>-1.8%</td>
</tr>
<tr>
<td>2015</td>
<td>2.2%</td>
<td>0.5%</td>
<td>-1.8%</td>
</tr>
<tr>
<td>2016</td>
<td>2.2%</td>
<td>0.4%</td>
<td>-1.8%</td>
</tr>
<tr>
<td>2017</td>
<td>2.1%</td>
<td>0.4%</td>
<td>-1.7%</td>
</tr>
<tr>
<td>2018</td>
<td>2.2%</td>
<td>0.5%</td>
<td>-1.7%</td>
</tr>
<tr>
<td>2019</td>
<td>2.3%</td>
<td>0.6%</td>
<td>-1.7%</td>
</tr>
<tr>
<td>2020</td>
<td>2.2%</td>
<td>0.5%</td>
<td>-1.7%</td>
</tr>
<tr>
<td>2021</td>
<td>2.0%</td>
<td>0.7%</td>
<td>-1.3%</td>
</tr>
</tbody>
</table>

*Note: Average annual growth rates of output per hour worked, UK, various Blue Book vintages. Source: Martin and Mackenzie (2021).*

US subsequently increased its productivity advantage over the UK until the early 1970s. The figure also shows that France and Germany’s labour productivity surpassed the UK by the mid-1960s.

The long-term trends of GDP per capita of the UK, US, France, and Germany provide a different picture of the UK’s economic performance evolution relative to other advanced economies (Figure 4). For instance, Germany’s GDP per capita is already higher than the UK by 1957. In addition, the UK’s GDP per capita caught up with the French GDP per capita again in 2002, and both countries had similar GDP per capita figures between 2002 and 2019.

These trends can be explained by the evolution of labour productivity (Figure 3) and that of hours worked per capita (Figure 5). Germany’s higher hours worked per capita allowed it to catch up with the UK’s GDP per capita as early as 1957, despite its lower levels of labour productivity at the time. On the other hand, France’s decreasing hours worked per capita since the mid-1970s allowed the UK to catch up with French GDP per capita in 2002 despite the UK economy remaining less productive than the French economy in terms of output per hours worked.

---

3This is consistent with the findings in Mason *et al.* (2018).
Figure 3: International comparison of labour productivity.

Note: Labour productivity in the UK, USA, Germany and France, expressed in USD 2010 PPP per hour. Source: Bergeaud et al. (2016).

2.3 An important measurement issue: hours worked

Aggregate labour supply can be explained by the employment rate (extensive margin) and the average number of hours worked per employee (intensive margin). Figure 6 contrasts the average number of hours worked by UK employees against averages in the US, France and Germany. The figure shows that the French and German workers, unlike their British and American counterparts, reduced their intensive margin of supply since 1970. The differences in terms of intensive labour supply are large enough to warrant a closer look at how these averages are measured. Marianna et al. (2018) find that countries, such as the UK, that use labour force surveys (LFS) data without further adjustments tend to overestimate the number of hours worked. To illustrate the size of this bias, the authors adjust the measures of hours using information available in the LFS and other complementary sources such as administrative or business statistics. They find that the hours worked measurement bias accounts for 10% or more of the relative labour productivity gap in many countries. For

4Employment rates can be decomposed into unemployment and participation rates.
instance, the US-UK productivity gap is reduced by a third when applying adjustments to the measurement of hours worked.

However, the country-specific hours worked adjustments introduced in Marianna et al. (2018) do not display disproportionate changes over time. This implies that while differences in hours worked measurements explain some of the international differences in productivity levels, they are less likely to explain international differences in labour productivity growth. Goldin et al. (2024) survey alternative sources of mismeasurement, including the changing boundaries of GDP and issues related to deflators’ measurement. The authors, however, do not report on the effects of mismeasuring hours worked.

2.4 Productivity growth accounting

The standard productivity growth accounting framework decomposes the growth of GVA per hour worked \((Y/N)\) into three components: labour composition \((Q)\), predominantly

Figure 4: International comparison of GDP per person.

Note: GDP per person in the UK, USA, Germany and France, expressed in USD 2010 PPP per person. Source: Bergeaud et al. (2016).
Figure 5: International comparison of hours worked per person.

Note: Hours worked per capita in the UK, USA, Germany and France. Source: Penn World Tables, version 10.0 (Feenstra et al., 2015).

representing the skills of the workforce; capital deepening ($K/N$), representing the availability of capital per hour worked; and a total factor productivity residual ($TFP$). The standard growth accounting framework is derived from the definition of TFP as a residual that captures changes in output unexplained by changes in production inputs (capital and labour)\(^5\)

\[
\Delta \ln(Y/N) = \Delta \ln TFP + \alpha \Delta \ln(K/N) + (1 - \alpha) \Delta \ln Q,
\]

where $Y$ represents output, $N$ hours, $Q$ a labour composition index, $K$ capital and $\alpha$ is the share of capital.\(^6\) Figure 7 presents the UK’s market sector cumulative gains in log labour productivity from 1970 to 2021 and its decomposition using the standard growth accounting method. The figure shows that 56% of the growth in the market sector output per hour between 1970 and 2019 is explained by improvements in TFP, about 34% can be attributed to capital deepening, measured as a capital over hours ratio. An improved composition

\(^5\)See appendix A for derivation details.
\(^6\)The capital shares $\alpha$ are also assumed to vary over time and are borrowed from ONS own growth accounting exercise.
Figure 6: International comparison of the average yearly hours worked per employee.

Note: Average hours worked per year and employee in the UK, US, Germany and France. Source: Penn World Tables, version 10.0 (Feenstra et al., 2015).

of the UK workforce from the mid-1990s onward culminates in contributing to 12% of the cumulative gains in output per hour by 2021.

Fernald et al. (2017) adopt a different approach to productivity growth accounting. They use a complementary decomposition of labour productivity in terms of the capital-output ratio

\[ \Delta \ln \left( \frac{Y}{N} \right) = \frac{1}{1-\alpha} \Delta \ln TFP + \frac{\alpha}{1-\alpha} \Delta \ln \left( \frac{K}{Y} \right) + \Delta \ln Q. \]  

(2)

This approach, which also derives from the definition of TFP, accounts better for the endogeneity of capital formation.\(^7\) In neo-classical growth models, slower TFP growth leads to slower capital deepening as measured using capital per hours worked (K/N) and no change in the capital-output ratio (Y/K). Changes to the capital-output ratio would then capture capital formation effects beyond what would occur endogenously due to changes in the TFP growth rate in the Solow-Swan model (Solow, 1956; Swan, 1956). Figure 8 shows the decomposition of the cumulative gains in the UK’s market sector labour productivity between

\(^7\)See appendix A for derivation details.
Figure 7: Standard growth accounting.

Note: Cumulative Decomposition of labour productivity from 1971 using standard growth accounting. Labour productivity is calculated as the market sector GVA per hour worked. ONS data, authors’ calculations (July 2022 data).

1970 and 2021 using the method in Fernald et al. (2017). Unlike the results from standard growth accounting, gains in labour productivity are almost exclusively driven by improvements in TFP until the mid-1990s when the effects of better labour composition start to be visible. The cumulative contribution of improved capital-output ratios remains close to zero throughout the studied period, with a slight positive effect for most of the great moderation period (1993-2004) and a small negative effect from the onset of the great recession until the start of the Covid 19 pandemic. The capital-output ratio improves following sudden drops in output, as it takes longer for capital to adjust. While the impact of the capital-output ratio on the cumulative gain in labour productivity since 1970 can turn from positive to negative and vice-versa, its overall contribution is never significant. This is indicative of a stable capital-output ratio in the long run.

The growth of the UK’s market sector labour productivity has slowed down substantially following the GFC. Table 3 shows that the average growth rate of output per hour in the
Figure 8: Fernald et al. (2017) growth accounting.

Table 3: UK growth accounting, two methodologies.

<table>
<thead>
<tr>
<th>Period</th>
<th>GVA/hour</th>
<th>Standard method</th>
<th>Fernald et al. (2017)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>lnY/N</td>
<td>lnTFP, α ln(K/N) (1 − α) ln Q</td>
<td>lnTFP, α ln(K/Y) ln Q</td>
</tr>
<tr>
<td>1997-2007</td>
<td>2.43</td>
<td>1.52, 0.62, 0.28</td>
<td>2.41, -0.43, 0.45</td>
</tr>
<tr>
<td>2009-2019</td>
<td>0.40</td>
<td>0.20, -0.13, 0.33</td>
<td>0.33, -0.45, 0.52</td>
</tr>
<tr>
<td>Slowdown</td>
<td>-2.03</td>
<td>-1.32, -0.75, 0.04</td>
<td>-2.08, -0.02, 0.07</td>
</tr>
</tbody>
</table>

Note: Average GVA per hour growth rate in the UK’s market sector in the 1997-2005 and 2009-2019 periods in % points, and its decomposition using the standard growth accounting method and the method in Fernald et al. (2017). The slowdown row represents the difference in the growth between the later and the earlier period. ONS data, authors’ calculations (July 2022 data).
market sector has dropped from 2.4% in the 1997-2007 period to 0.4% in the 2009-2019 period. The standard growth accounting and the Fernald et al. (2017) method draw different pictures when decomposing the post-GFC productivity slowdown. Standard growth accounting explains about a third of the slowdown by capital shallowing while the capital-output method attributes nearly all the slowdown to weaker TFP growth. Both methods show that the improvements in labour compositions since 2009 marginally attenuated the post-GFC productivity slowdown. We are thus drawn to an examination of why the growth in TFP and capital accumulation has disappointed so.

3 The recent productivity slowdown under the neoclassical lens

We put the recent UK productivity slowdown under the lens of the Solow-Swan standard growth framework and show that this framework cannot predict the rise in hours worked after 2008. Higher labour since 2008 was not however accompanied by more investment as would be suggested by neoclassical economic theory. We then attempt to reconcile the slowdown of TFP and the lower real interest rates using the capital accumulation neoclassical growth model. We find that the drop in real interest rates after 2008 is only consistent with post-GFC TFP slowdown for extreme values of the households’ elasticity of intertemporal substitution.

3.1 The Solow-Swan model and long-term trends

The Solow-Swan model (Solow, 1956; Swan, 1956) is the standard literature framework for studying long-term economic growth. The model relies on exogenous total factor productivity (TFP), saving rate and population growth to explain the long-term trends of other macroeconomic variables. In the Solow-Swan model, production is constrained by a Cobb-Douglas technology using two inputs: labour hours and capital; TFP represents the residual output unexplained by changes in the two production inputs.

8 The results emanating from the standard decomposition are consistent with other studies providing an accounting of the labour productivity slowdown in the UK using ONS data (e.g., Riley et al., 2018; Goodridge et al., 2018).

9 Goldin et al. (2024) use EU KLEMS instead of ONS data and find a more substantial contribution of labour composition to the slowdown of labour productivity in the United Kingdom. As noted by the authors, the EU KLEMS implied labour composition index differs across vintages, hence our preference for the more consistent ONS data.
In addition to hours and capital, modern accounting frameworks consider the impact of labour composition (or labour quality) on production. We, therefore, modify the Cobb-Douglas production function used in the Solow-Swan model to account for the labour composition index $Q$

$$Y = TFP \times K^\alpha (Q.H)^{1-\alpha},$$ (3)

where $Y$ is output, $K$ is capital, $H$ hours, and $\alpha$ is the share of capital. In the modified Solow-Swan framework, long-term growth of labour, capital, and hours are

$$g(Y) = g(K) = \frac{g(TFP)}{1-\alpha} + g(Q) + g(N),$$ (4)

$$g(H) = g(N),$$ (5)

where the function $g$ is used to designate the long-term growth rate and $N$ is the population size. These are textbook Solow-Swan long-term growth results (e.g., Romer, 2012), with the additional contribution of labour quality to the long-term trends of output and capital.

Table 4 shows the average growth rates of TFP, labour composition, working-age population, hours worked, output, capital and the average growth rate predicted by the model for hours, capital and output for the pre-GFC period (1990-2007) and the post-GFC period (2008-2019). Except for the average growth of hours worked, Solow-Swan trends account well for the pre-GFC long-term trends. Over the 1990-2007 period, the average growth rate of hours worked (0.17%) is about a third of the average growth rate of the UK’s working-age population (0.5%). Given the long-term growth rates of the working-age population, labour composition and TFP, the model predicts long-term growth of capital and output of around 2.69% between 1990 and 2007. Historical data show that output grew on average at a rate slightly below the one expected by the model (2.46%) and that capital grew at a rate close to the one expected by the model (2.6%).

In the post-GFC period, hours worked grew at an average of 1.07% per year, significantly outpacing the growth of the working-age population (0.34%). On the other hand, capital grew at a rate close to the one expected by the model (0.87% in the data and a model-implied growth of around 1%). The discrepancy between the model-implied growth of output and the much higher average output growth in the data (1.41%) is mainly explained by the unusual growth of hours worked in the decade that followed the global financial crisis.

Admittedly, labour supply is exogenous by design in the Solow-Swan model so this neoclassical growth framework can hardly be criticised for missing the changing trends in the UK’s labour market. Textbook expositions of the model assume that labour supply is inelastic and driven by population growth. This is also the main interpretation of the exogenous
Table 4: Average growth rate of UK macro variables and the prediction of the Solow-Swan model.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Solow-Swan model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TFP</td>
<td>Q</td>
</tr>
<tr>
<td>Pre-GFC</td>
<td>1.08</td>
<td>0.44</td>
</tr>
<tr>
<td>Post-GFC</td>
<td>0.12</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Note: Average growth of macroeconomic variables and the Solow model’s predicted output and capital average growth rates expressed in percentage terms. The pre-GFC period covers from 1990 to 2007 and the post-GFC period covers 2008 to 2019. The model predictions in both periods are made using the period’s average growth rates of TFP, labour quality $Q$, and working-age population $N$ to predict the average growth of hours worked $H$, output $Y$ and capital $K$. ONS data for the UK’s market sector economy and the UK’s population aged between 16 and 63, authors’ calculations (ONS data from July 2022).

labour supply in Solow (1956). We use the working-age population as a slightly better demographic proxy of aggregate labour supply.\footnote{Fernández-Villaverde et al. (2023) argue that considering GDP per working-age adult instead of GDP per capita explains away much of the apparent weak performance of the Japanese economy relative to the United States.} Within the Solow-Swan framework, the rapid rise in hours post-2008 should have led to an increase in the marginal product of capital, thus leading to substantially more investment than recorded after the great financial crisis. As noted above, capital growth in the post-GFC period was in line with the Solow prediction, while hours grew faster than expected by the Solow-Swan model. This means that the Solow-Swan model cannot reconcile investment levels with the growth of aggregate hours worked.

In summary, the Solow-Swan model accounts well for the long-term trends of capital, labour hours and output in the pre-GFC period. Given that it assumes labour supply to be exogenous, the Solow-Swan framework is unable to predict the post-GFC increase in aggregate hours. Crucially, while this growth framework predicts a constant capital-output ratio, post-GFC capital grows at a slower pace than output. The lack of investment despite abundant labour supply and cheap financing costs after 2008 is a puzzle that evades the explanatory power of standard neoclassical growth theory.

### 3.2 Long-term interest rates trends under a neoclassical lens

The Ramsey rule is a good starting point for discussing long-term saving behaviour under a neoclassical lens. The rule stipulates that the marginal product of capital $r$ is a function of the rate of time preferences $\rho$, the expected real growth rate of per-capita consumption $g(c)$...
and the elasticity of marginal utility of consumption $\sigma$ (hereafter called EMU)

$$r = \rho + \sigma g(c).$$

(6)

The original contribution by Ramsey (1928), was developed further by Cass (1965) and Koopmans (1965), leading to the now standard Ramsey-Cass-Koopmans growth model (hereafter RCK). The model assumes that households’ utility is a function of the per-capita consumption $c$

$$U = \frac{c^{1-\sigma}}{1-\sigma},$$

(7)

where $\sigma$ is the same EMU that appears in the Ramsey rule and $1/\sigma$ corresponds to the elasticity of intertemporal substitution (EIS). In addition, the RCK model also assumes that the representative households seek to maximise their lifetime utility, discounted at a rate $\rho$. Similarly to the previous subsection on the Solow-Swan model, we modify the Cobb-Douglas production assumed in the standard version of the RCK model to allow for changes in labour quality as in expression (3). As in the Solow-Swan model described above, the growth rate of consumption per head is

$$g(c) = g(TFP) \frac{1-\alpha}{1-\alpha} + g(Q).$$

(8)

The Ramsey rule is then derived from the household’s optimal behaviour and can be written as

$$r = \rho + \sigma \left\{ g(TFP) \frac{1-\alpha}{1-\alpha} + g(Q) \right\}.$$

(9)

Notwithstanding a change in the model parameters $\rho$ and $\sigma$, the modified RCK growth model predicts that, as TFP growth slows down following the global financial crisis (GFC), real rates drop by $-\Delta r$

$$\Delta r = \sigma \left[ g(TFP_2) \frac{1-\alpha_2}{1-\alpha_2} - g(TFP_1) \frac{1-\alpha_1}{1-\alpha_1} + g(Q_2) - g(Q_1) \right],$$

(10)

where $TFP_2$ is the post-GFC long-term growth rate, $Q_2$ is the post-GFC labour composition index growth rate, $\alpha_2$ is the post-GFC capital share and subscript 1 is used for the pre-GFC quantities.

The standard business cycle model enables the modelling of economic fluctuations at the business cycle frequencies (Kydland and Prescott, 1982; Long and Plosser, 1983). These models can be overlaid on top of the discrete-time version of the RCK model, thus generating

\footnote{For a more detailed treatment of the continuous-time version of the Ramsey-Cass-Koopmans growth model, see chapter 2 of Romer (2012). For a treatment of the discrete-time version of the model, see Novales et al. (2009).}
both the RCK long-term trends and the business cycle fluctuations.\textsuperscript{12} However, the business cycle fluctuations vanish in the long run and only the trends remain. Future households’ utility is discounted using a factor $\beta$ that depends on the rate of time preferences $\rho$ and TFP growth rate $g(TFP)$

$$\beta = e^{-\rho-\sigma \left( \frac{g(TFP)}{1-\alpha} + g(Q) \right)}.$$ \hspace{1cm} (11)

A slowdown in TFP growth leads to an increase in the discounting factor $\beta$.\textsuperscript{13}

The Ramsey rule underpins the determination of the social discount rate (SDR) crucial to the social welfare function used by many government agencies in their cost-benefit analysis of long-term projects. This is the practice in the UK (HMT, 2023), France (Lebègue, 2005) and many other countries. Moreover, the Ramsey rule is consistent with the integrated assessment models used in climate change economics (Nordhaus, 2014).

From its inception, much of the debate around the Ramsey rule focused on the rate of pure time preferences $\rho$. While providing the theoretical basis for a positive $\rho$, Ramsey argued that “we do not discount later enjoyments in comparison with earlier ones, a practice which is ethically indefensible and arises merely from the weakness of the imagination” (Ramsey, 1928). In Ramsey’s altruistic tradition, Stern (2006) adopted a pure rate of time preferences of $\rho = 0.1\%$, solely justified by “[the] uncertainty about [the] existence of future generations arising from some possible shock which is exogenous to the issues and choices under examination”. Nordhaus (2014), on the other hand, assumes a higher rate of time preference at $\rho = 1.5\%$, thus heavily discounting the long-term benefits (beyond a century) of climate change investment.

Data from the Bank of England show that the 1-year real rates dropped from an average of around 4.1\% in the 1970-2007 period to an average of -1.3\% between 2008 and 2016, which is a drop of 5.4\%.\textsuperscript{14} Notwithstanding a change in the rate of pure time preferences $\rho$ and the EMU $\sigma$, expression (10) can be used to understand the drop in real interest rates. The RCK model-implied interest rate drop would depend on the assumed value of $\sigma$. Borrowing the EMU value from the UK’s Green Book ($\sigma = 1$) the model implies that interest rates are 1.5\% lower in the post-GFC period. Best \textit{et al.} (2020) survey the literature on estimating the UK’s EMU and find significant heterogeneity in the values of $\sigma$. The author’s retained central value ($\sigma = 1.5$) implies a post-2008 interest rate drop of 2.3\%. Best \textit{et al.} (2020) exploit a quasi-experimental variation in the UK mortgage interest rates to estimate a small

\textsuperscript{12}This exercise, however, requires non-trivial conditions on both agents’ preferences and the technology constraining production (King \textit{et al.}, 1988).
\textsuperscript{13}This guarantees that the return on capital, derived from the business cycle standard steady-state Euler saving condition, yields the same return on capital as the RCK model $r = 1/\beta - 1 \approx \rho + \sigma \left( \frac{g(TFP)}{1-\alpha} + g(Q) \right)$.
\textsuperscript{14}See Figure 20 below.
EIS of around 0.12 for the UK (or an EMU of around 8), thus implying a post-2008 drop in interest rates that can hardly be reconciled with the data (12%).

Conversely, one can use expression (10) to estimate the UK’s EMU. Referring to the 1-year real interest, the RCK model implies an EMU of 3.5. On the other hand, if one uses the drop in the loan rate facing safer UK non-financial corporations\textsuperscript{15}, the RCK model implies the EMU value $\sigma = 2.3$. Naturally, considering the higher interest rate environment post-2021 would moderate the drop in real rates, thus leading to lower model-implied EMUs for the UK. One can refer to the notion of natural real interest rates to do away with the effects of business cycle fluctuations. Holston et al. (2017) estimates point to a 1.1% drop in the natural interest rate between 2017 and 2017, implying an EMU around 0.7. Rachel and Summers (2019) document a drop in $R^*$ of about 50bp in advanced economies following the GFC. This would imply a much lower value of $\sigma$ around 0.3.\textsuperscript{16}

A potential secular stagnation in productivity growth is likely to impact how economists discount the future. The change of discounting can happen because of a lower consumption growth rate, a change in estimates of the EMU as argued above, or increased uncertainty over future consumption trends.\textsuperscript{17} Given the nature of the UK productivity slowdown all three factors will likely lower discount rates. Nonetheless, debates on discounting are far from settled and given their relevance to public investment and the green transition policies, they are likely to persist and attract further research effort.

4 The UK demographic trends and productivity

This section reviews the ‘secular stagnation’ debate sparked by the inability of advanced economies to bounce back following the global financial crisis, focusing on the ramifications of population ageing on economic growth. We then examine the UK’s demographics of labour supply post-2008 and ask questions as to how the UK’s specific labour supply dynamics may have affected productivity growth.

4.1 The revival of the secular stagnation debate

The question of the impact of demographic trends on economic growth is not new. Hansen (1939) coined the concept of ‘secular stagnation’ in his study of the possible long-term slowdown of the US economy due to a reduction in population growth in the 1930s. Hansen,\textsuperscript{15} See Figure 21 below. The figure shows that the long-term average of real loan rates facing safer non-financial UK corporates dropped from 3.5% to about 0%.

\textsuperscript{16} The estimates of $R^*$ in Holston et al. (2017) and Rachel and Summers (2019) both use the methodology introduced by Laubach and Williams (2003).

\textsuperscript{17} Gollier (2012) corrects the Ramsey rule to allow for uncertainty in the growth rate of per-capita consumption is \emph{i.i.d} normal with mean $\bar{g}$ and variance $\text{Var}(g(c))$: $r = \rho + \sigma \bar{g} - 0.5 \sigma (1 + \sigma) \text{Var}(g(c))$.\textsuperscript{16}
in turn, relied on the much older ideas put forward by Adam Smith regarding the positive
effects of population growth in fostering demand and improving the production of new ideas.  
The original argument by Hansen (1939) was that slower population growth would depress
real interest rates as capital will be combined with less labour input at the margin. However,
this argument ignores the observation that aging populations with higher dependency ratios
will decrease aggregate savings. On this basis, Goodhart and Erfurth (2014) predict that
real interest rates will return to their historical equilibrium value of around 2.5-3%, as did
Chadha and Dimsdale (1999) in earlier work.

Larry Summers brought the term ‘secular stagnation’ back to life in 2014 (Summers,
2014). Summers noted (i) the persistent negative gap between measures of potential and
actual output in most advanced economies and that (ii) US and world real interest rates have
declined since the mid-1980s. He concluded that this indicated a much lower full employment
real interest rate (FERIR). He argued that the nominal interest rates’ zero lower bound and
the low inflation environment prevent the real interest rate from dropping low enough to
meet the FERIR required to bring unemployment down.

Rachel and Summers (2019) point to the high capital mobility between advanced
economies and the limited fluctuations in their aggregated current account. On this basis,
they argue that the natural real interest rate R* is better estimated for an economic bloc of
all advanced economies. The authors document a decrease of R* by 300bp in the advanced
economies bloc between 1970 and 2017 and argue that but for extraordinary fiscal policy,
the natural real interest rate may have fallen by an extra 400bp. They argue that the
increases in government spending have obscured the substantial shifts in private saving and
investment propensities since the early 1970s.

Estimates of potential output have been revised downward as economists started to factor
in the lower productivity growth regimes in the US and most advanced economies after
2008 (e.g., Gordon (2014); Kahn and Rich (2007) and (2022); Fernald (2015)). Figure 9
shows the series of downward revisions to productivity trend forecasts by the UK’s Office
for Budget Responsibility. The repeated downward revisions to productivity trends and
potential output measures moved the debate from the issue of a large negative output gap, as
in Summers (2014), to the anaemic growth in potential output as productivity growth slows
down (Gordon, 2014). Discerning new productivity growth trends from cyclical fluctuations
is crucial for short-term economic policy. For instance, many economists argue that the
failure to detect the productivity slowdown after the 1973 oil shock contributed significantly
to the economic instability of the 1970s (e.g., Kahn and Rich, 2007). More recently, Philippon
(2022) argues that the assumption of exponential TFP growth contributed to overestimating
productivity growth, thus overestimating potential output after 2008. Philippon suggests
that a linear TFP growth model is more suitable for forecasting productivity. If linear TFP
Figure 9: Vintages of UK and US trend productivity forecasts.

Note: UK trend productivity is defined as potential non-oil GVA over potential hours, while US trend productivity is defined as the potential non-farm business output divided by potential hours worked in the non-farm business sector. Source: OBR (2016).

Growth is indeed the way forward, the current practice of assuming exponential TFP growth would lead to further overestimation of productivity trends and potential output. Greenwood et al. (1997), Fisher (2006) and others argue that the relative prices of investment goods have been trending lower since the early 1950s, with a notable acceleration in the rate of decline in the early 1980s.\(^{18}\) IMF (2014) finds that the trend of decreasing investment prices extends to many other advanced and emerging economies. The increase in savings supply from emerging economies and the drop in capital demand due to cheaper capital goods are then presented as a potential justification for the secular decrease in real rates after 1980.

Both Gordon (2015) and Fernald (2015) explain the productivity slowdown as the end of an exceptional productivity growth period that started in the mid-1990s with the improvements in the IT producing and using industries. Fernald (2015) dates the slowdown in US TFP to circa 2005, i.e., several years before the onset of the great recession. Focusing on the UK productivity puzzle, Fernald and Inklaar (2022) argue that most of the UK’s TFP slowdown follows from the slowdown in the US frontier. There is a view that productivity improvements do not happen at a steady, even pace. Instead, productivity improves much faster in some eras (Gordon, 2016). According to this view, some inventions are more important than others; while improvements in IT-related technologies were responsible for the most recent episode of productivity revival (1995-2005), earlier inventions such as electricity

and the combustion engine were far more favourable to productivity growth as they could sustain productivity gains in advanced economies for a century (1870-1970).

The effects of demographic dynamics on macroeconomic variables have been subject to intense debates since the re-emergence of secular stagnation as an important line of inquiry. On the one hand, older segments of the population tend to save less and this contributes to increasing real interest rates. On the other hand, longer life expectancy makes the active part of the population keener on saving to fund consumption for extended retirement periods, pushing real interest rates lower. There is also no agreement in the literature on the effects of ageing in advanced economies on economic growth. Acemoglu and Restrepo (2017) argue that ageing societies are more likely to adopt new automation technologies, thus increasing GDP per capita, while Aksoy et al. (2019) show that ageing decreases innovation, causing slower economic growth and less investment, thus depressing real interest rates. Consistent with the idea that population ageing hinders innovation, Jones (2022) uses standard growth models with exogenous population growth and a model with endogenous fertility to show that knowledge and living standards stagnate when demographic growth turns negative. In a recent paper, Hopenhayn et al. (2022) argue that a decrease in population growth can reduce firm entry, thus affecting firm dynamics in a way that increases industry concentration, reduces the labour share and harms productivity growth.

4.2 The demographics of UK labour supply

The UK labour market reforms in the last three decades of the twentieth century created the conditions for a flexible response to the recession. Tax reforms were used to shift employers’ incentives and unemployment (and non-participation) was reduced. There was also an increase in the labour market’s flexibility with some trade union powers and employment protection legislation reform. Trade union power also diminished because of the decline in traditional manufacturing and public sector monopolies. In effect, we may have stumbled on a low wage, low productivity and high employment equilibrium. This outcome limited the impact on unemployment from the Great Recession and may have played a role in reducing the extent to which households increase their savings ratios. The maintenance of relatively elevated employment levels may have helped limit the impact on house prices from the recession and thus limited spillovers to a vulnerable financial sector. The other side of this high employment outcome might be an anaemic productivity growth path.

Taking a closer look at the demographics of labour supply in the UK in the post-GFC period, two clear trends emerge. First, the average hours worked remained relatively stable since 2008 (Figure 10), making the increase in aggregate hours a consequence of higher participation and low unemployment rates. In other words, there was no noticeable change
in labour supply at the intensive margin, with a marked increase in supply at the extensive margin. In addition, the increase in aggregate labour hours is primarily driven by higher employment rates among older workers (50 and older), especially older women. This is clearly displayed in Figure 11. The figure shows that the employment rate for the 50-64 age group increased from 65% in 2007 to 72% in 2019. In contrast, the employment rate of the 35-49 age group has only increased by 3% in the same period. Decomposing the growth of aggregate labour hours by age group, an ONS study (Syed et al., 2019) shows that the drop in hours worked during the post-GFC recession was concentrated in the younger workers demographic (16 to 29 years old). After the end of the recession, employment hours among young workers remained low while the work hours supplied by older workers soared (Figure 12).

The increase in labour supply by older women gives some credence to the flexible labour supply explanation, as the increase coincides with changes in the State Pension age (SPA) after 2007 that pushed the retirement age for women from 60 to 65. Following these changes, women delayed their permanent exit from the labour market by an average of two years, with the average exit age increasing from 62 in 2007 to over 64 in 2019 (Figure 13).
Another remarkable change in the composition of the UK workforce is the marked increase in measures of workers’ skills. The share of hours supplied by workers with a university degree increased from 25% in 2008 to 35% in 2018 (Syed et al., 2019).

In summary, the increase in aggregate hours after 2008 has been driven by older workers, especially older women. This trend is not simply reflective of the ageing of the UK population. It also reflects higher employment rates among older workers and a decrease in employment among younger workers. Experienced workers are usually assumed to possess more human capital, so this compositional change deepens the UK’s productivity slowdown puzzle. Similarly, the productivity puzzle is deepened by the increase of the portion of workers with higher education qualifications.

The role of the flexible labour market in increasing aggregate hours after 2008 has been a subject of interest in the literature. For instance, Pessoa and Van Reenen (2014) note that wages, deflated by the GDP deflator, fell by 4% in the four years following Q2 2008. This drop in real wages, they argue, is unprecedented for a post-war recession, reflecting the effect of a long-term policy in favour of flexible labour supply in the UK (weaker unions;
work search pressure on benefit claimants, etc.). They attribute low post-2008 investment to high capital costs and the uncertainty plaguing the UK economy. However, the latter explanation for weak investment is more likely, considering the data regarding the cost of capital after 2014 and the additional uncertainty brought about by Brexit (Sampson, 2017).

This argument gains further support from Douch et al. (2023), who suggest that the abundant labour may have contributed to the underwhelming productivity performance in the UK. They focus on the cohort effect of firms established after 2008, showing a post-GFC redeployment of labour from the slowing high-productivity financial services industry towards lower productivity service industries. This phenomenon, which they dub a productivity “levelling down”, highlights the impact of labour market flexibility on productivity trends.

An alternative explanation for the higher labour supply by the UK workforce lies in the increase of households’ debt during the credit expansion that preceded the global financial crisis. Bunn et al. (2021) exploit data from the Labour Force Survey (LFS) and the Wealth and Assets Survey (WAS) to show that a negative shock to income leads to a reduction in labour force participation among outright homeowners while increasing the desired hours...
of mortgage holders. In addition, the authors find that households with higher debt levels increase their labour supply by more following shocks that decrease their ability to service debt (adverse income shocks; higher interest rates). The authors then argue that the increase in UK households' debt in the years leading to the financial crisis has the potential to explain, at least partly, the subsequent behaviour of employment, hours and wages.

Post-pandemic ONS data show increased labour market tightening and an increase in unfilled job vacancies (ONS, 2021). Employment dropped by 1.4% in the year ending September 2021, primarily driven by younger workers and elementary occupations. It remains unclear if the post-pandemic labour trends will fade as the economy recovers from the successive shocks of Brexit and Covid-19 or if they represent permanent new trends triggered by the pandemic shock.

Both the flexibility of the labour market and households’ indebtedness are likely to have contributed to the increase in aggregate hours supplied by the UK workforce in the post-GFC decade. However, important questions remain. As noted above, the compositional changes to the UK workforce deepen the country’s productivity puzzle. Moreover, the increase in labour supply was not accompanied by an increase in investment, as economic theory would predict. The following subsection takes a closer look at the investment trends in the UK.
Figure 14: Gross fixed capital formation shares of GDP, an international comparison.

Note: Gross fixed capital formation to GDP in current prices in the UK, US, France and Germany. Prior to 1991, West German data is used for Germany. Source: Eurostat (2022 data).

4.3 Investment trends in the UK

The UK economy suffers from chronic underinvestment. Figure 14 shows that the UK’s investment share of GDP has been significantly lower than the ratios in France, Germany, and the US for most years since 1965 and continuously since the early 1990s. The share of GDP dedicated to business investment has trended downward since 1965 (Figure 15), while the long-term average net public investment fell from 4.5% of GDP between 1949 and 1978 to 1.5% from 1979 to 2019 (Figure 16). As a result of these weak investment trends, the net capital stock to GDP ratio trended downward after 1960 (Figure 17).

There are structural reasons for these trends. For instance, capital goods’ prices decreased substantially relative to other goods over the last half-century. Figure 18 shows that the relative price of investment trended downward in the UK after 1990, with improved ICT industries and computing power often cited as an important factor behind lower relative investment prices. Leon-Ledesma and Moro (2020) build a two-sector structural change model with a high productivity growth good sector and a low productivity growth services sector.
and show that as the share of services in the economy increases, the relative price of investment decreases, thus accelerating investment-specific technological change. The authors assume a unitary capital elasticity of labour, which implies that the nominal investment share of GDP remains constant in their model while the real investment share increases.\footnote{Leon-Ledesma and Moro (2020) calibrate their model to the United States, where the nominal share of investment remains approximately stable in the period they study. This is consistent with the behaviour of the US nominal investment share in Figure 14.} Thwaites (2015) documents the drop in the relative price of investment in several advanced economies before building an overlapping generations model parameterised with a less-than-unit elasticity of substitution between labour and capital. The low substitutability of labour for capital means that the economy reacts to lower relative prices of capital by reducing the share of investment in nominal GDP. As shown in Figure 18, other advanced economies witnessed similar, if not larger, drops in the relative investment price to the UK after 1991. Moreover, Podkaminer (2019) finds no evidence of the low relative price of investment goods Granger-causing the decline of investment shares in industrial countries. This means that the UK’s secular underinvestment trends remain a puzzle.

The UK’s withdrawal from the European Union and its trade consequences have increased uncertainty for UK businesses. This uncertainty is likely to have contributed to reversing the
short-lived gains in the GDP share of business investment between 2014 and 2016 (Figure 15). Górnicka (2018) uses firm-level data to show that potential trade costs have had a considerable and statistically significant negative impact on firm investment in the UK after the referendum. The pandemic-related economic disruptions compounded the terms of trade uncertainties and are likely to have further depressed business investment (Figures 15 and 19).

Another feature of the weak investment puzzle in the UK is that it is happening in the context of low real interest rates. This deepens the weak investment puzzle in the UK. The following section looks at this issue in more detail by considering the UK’s financial landscape and how it might impact investment.

5 Finance and productivity

The link between finance and productivity is complex. It is unclear whether finance ignites productivity or flows from expected future productivity gains, making causal identification difficult to achieve. To deal with these endogeneity issues, Butler and Cornaggia (2011)
exploit an exogenous shift in demand for corn to study how the depth of the local banking sector impacts farmers’ productivity. The results indicate that productivity increases more in areas with better financial intermediation as measured by aggregate deposits in local banks.

Consistent with the causal finance to productivity link documented in Butler and Cornaggia (2011), Levine and Warusawitharana (2021) use firm-level data to show that an increase in financial frictions post-2008 led to a higher sensitivity of productivity growth to the availability of external finance.\(^\text{20}\)

Accordingly, this section examines the UK business and financial landscape, documenting the puzzlingly low business investment trends despite the decline in real interest rates. We then turn to theoretical explanations of the low real rates-low investment puzzle, starting with the zombie firms’ explanation before considering more recent theoretical developments.

\(^{20}\)The authors deal with endogeneity issues by decomposing TFP into expected and unexpected components (as in Levinsohn and Petrin (2003)).
Following the period of volatile credit markets and a tight supply of credit associated with the
global financial crisis (GFC), the Bank of England reacted by bringing its main interest rate
to 0.5% in March 2009, down from 5.25% a year earlier. Low bank rates were accompanied
by a large asset purchase program, as the Bank of England accumulated £895 billion in UK
government bonds and £20 billion in UK corporate bonds by the end of 2020. While the
Bank of England’s easing pushed real interest rates into negative territory post-2008, the
downward trend in real interest rates predates the global financial crisis. Figure 20 shows
that real interest rates have trended down since the early 1990s. After the global financial
crisis, the real interest rates facing safer non-financial private companies have collapsed from
a long-term average of 3.6% to a post-GFC average near zero (Figure 23).

Against this backdrop of low real interest rates, investment, particularly business investment,
remained sluggish for most of the decade following the GFC. Figure 19 shows that
Figure 19: UK real business investment.

Note: UK real business investment index (1997 Q1 = 100). Source: ONS (February 2022 data).

Figure 20: UK real interest rates.

Note: UK real rates interest rates. Source: Bank of England (Thomas and Dimsdale, 2017)
it took eight years for the UK’s real business investment to reach its 2007 level. This recovery was short-lived, however, as business investment stalled in 2016 following the Brexit referendum result and reversed with the start of the pandemic-related economic disruptions.

Despite the abundant liquidity and low real interest rates, small and medium enterprises (SMEs) continued to be credit-constrained, pushing them to rely on more expensive forms of credit. Brown et al. (2019) report that in the fourth quarter of 2015, 17.1% of UK SMEs used credit card financing and that a similar portion of SMEs used bank overdrafts (17.4%). Only 7.7% of UK SMEs used the typically much cheaper bank loans and commercial mortgages in Q4 2015, indicating that the SME sector may face severe credit constraints. The behaviour of credit markets during the Covid-related economic disturbances is another indication of the credit constraints facing SMEs. The UK government launched multiple loan guarantee programs to ease the credit conditions in the context of weaker business cash flows caused by the pandemic-related restrictions. The UK’s private non-financial corporations raised around £75.5 billion of debt between March 2020 and May 2020. Almost all of this figure
Figure 22: The financial sector is concentrated in London.

Note: Greater London’s Banking and Insurance sector GVA to the UK’s overall GVA of the sector in Current Prices. Source: ONS, authors’ calculations (June 2021 data).

(£75 billion) was raised through the government’s Covid-19 lending schemes. Of the £75 billion borrowed through these schemes, only £5.6 billion is estimated to have gone to larger businesses (Samiri, 2022). The important loan take-up by SMEs through the government’s Covid guarantee schemes indicates unsatisfied financing demand within the SME sector. The European Investment Fund’s SME Access to Finance Index shows that the United Kingdom underperformers competing EU economies such as Germany and France in supplying finance to SMEs (Torfs, 2021). Subcomponents of the SME Access to Finance index show that UK SMEs struggle to access loan financing while being better served in the equity finance markets.

The UK financial system is concentrated in London and a few secondary cities. Wójcik and MacDonald-Korth (2015) find that the UK financial system concentrated further between 2008 and 2012, with London’s share of the UK’s financial employment moving from 31% to 34% in the four years period following the global financial crisis. Figure 22 shows that the concentration of the financial industry in London continued after 2012. By 2019, over half of the UK’s Banking and Insurance sector’s GVA came from the Great London area (as opposed to 48% in 2012 and 46% in 2007).
Figure 23: Regional GVA and the SME loan financing.

Given the uneven geography of equity finance, particularly that of venture capital (Mason and Pierrakis, 2013), it is likely that SMEs in peripheral regions are more dependent on loan financing. Figure 23 shows that SMEs outside London and the England South East had a higher bank debt financing to turnover ratio in 2016. This figure, however, fails to disentangle the demand-side from the supply-side effects. Using firm-level data from Q1-2011 to Q3-2013, Lee et al. (2015) found that SMEs in the 10% least accessible NUTS2 regions (peripheral regions) applied for more loans than SMEs in the non-peripheral areas. The authors also find that innovative SMEs in remote regions suffer from a 'liability of distance' when attempting to access funding. After controlling for financial history and other financial risk indicators, the authors find that these SMEs are more discouraged from applying for bank loans and face higher rejection rates. However, the authors find no significant adverse geographic effect impacting the access to bank loans of non-innovating firms located in peripheral regions.

Uncertainty, resulting from significant news, might delay and defer financial investment, thus hampering improvements in human and physical capital. Capital allocation may be quite impressively allocated from the City of London to the rest of the world, but this does...
not necessarily mean that capital finds its way to small and medium-sized enterprises outside London and the South East. Daams et al. (2023) use commercial real estate investment data to show that the GFC triggered a flight to safety of capital into London, thus deepening the UK’s regional inequalities.

The UK’s SME funding gap is an old question first tackled by the Macmillan Committee in 1931. One study estimates this gap at around £22 billion (NAO, 2013). More recent studies argue that an SME regional funding gap is concentrated in equity financing (Stansbury et al., 2023; Oliver Wyman, 2024). This debate, however, remains unsettled and is likely to attract more research effort in the future.

In the remainder of this section, we draw on the economic literature to describe several theoretical mechanisms that help shed light on the reasons behind the UK’s weak productivity growth in the context of the financial landscape described above.

5.2 Banks and zombie firms

Since the global financial crisis, banks have undergone a process of repairing their balance sheets and building up capital and liquidity buffers. In the immediate aftermath of the crisis, there had been a relatively low level of insolvencies, given the depth of the recession. The argument has been made that the forbearance shown by banks towards existing firms and the lack of availability of finance to new firms had acted to reduce the introduction of new technologies into the overall production function. Arrowsmith et al. (2013) found that forbearance might account for some of the shortfall in productivity, but as they only measure the impact on SMEs, the whole economy impact may be considerably larger. In a Euro Area analysis, Andrews and Petroulakis (2019) suggest that relatively undercapitalised banks were more likely to show forbearance and may have played a role in starving credit to healthier firms. Acharya et al. (2019) argue that while Outright Monetary Transactions helped re-establish stability in the banking sector, they did not fully translate into economic growth. The authors explain that banks were more likely to extend loans to distressed firms that built up their cash reserves rather than invested, thus misallocating funding in sectors with a high prevalence of zombie firms.

Firms that have tended to be unprofitable over time and have a low stock market valuation tend to invest less and be more vulnerable to shocks. Using low market stock valuations and low-profit rates to identify zombie firms, Banerjee and Hofmann (2022) document the rise of their share from some 4% in the 1980s to 15% by 2017 across advanced economies and over 20% in the UK. In earlier work, Banerjee and Hofmann (2018) present evidence of a positive link between low rates and the number of zombie companies at the country and sectoral

---

\(^{21}\)This figure of £22 billion has been used by others (e.g., Carney, 2019).
levels. That said, Favara et al. (2021) do not think the zombie firm issue is that significant in the US and accounts for a “small share of credit to non-financial firms”. But Gouveia and Osterhold (2018), who focus on Portugal, argue strongly that the prevalence of Zombie firms has dragged down aggregate productivity and their tendency to stay in the market affects intra-sectoral productivity. Extending this analysis to the set of OECD economies, McGowan, Andrews and Millot (2017) document that an increasing share of industry capital has become attached to zombie firms. This has tended to limit the expansion of healthy firms and led to market congestion, which heightens barriers to entry. They argue that the rise of such firms has affected potential output growth, investment and the accrual of TFP.

More work is required to obtain consensus on a working definition of zombie firms in the UK. This would help in documenting further the zombie firms’ issue in the UK before estimating its effects on aggregate productivity growth.

5.3 Theoretical arguments for low productivity growth in a low interest rate, ample liquidity environment

Real interest rates are at a historic low (Figures 20 and 21), yet business investment did not soar following the GFC. This puzzle is difficult to explain in the context of standard macroeconomic models that usually focus on frictions at the financial intermediation level. For instance, Stiglitz and Weiss (1981) argue that, under imperfect information, interest rates can be used by the lending bank to sort borrowers. The crux of the argument is that only risky borrowers are more likely to take loans at a high interest rate (selection effect). To increase the average quality of its loan portfolio, the representative bank would then avoid supplying credit at an interest rate that it deems high enough to attract bad borrowers, thus preventing the credit market from clearing and rationing credit for the riskiest subgroup of borrowers. However, it is unclear if this and other similar credit rationing models are consistent with low aggregate investment in a low interest rates environment.

A new generation of models focuses on the firm’s behaviour when financing is cheap. Liu et al. (2022) describe a mechanism where low interest rates can lead to more market power for industry leaders, thus stifling innovation and productivity gains in the long run. When interest rates are low, the safe profits emanating from a permanent monopoly position are discounted at a lower rate. This increases the value of the permanent monopoly position from the perspective of a market leader. The market leader is then incentivised to invest in innovation to capture the permanent monopoly profits. The market follower can also invest in research and development (R&D) to improve its competitive position. However, in the case of successful innovation, the market follower would enter a fierce ’neck to neck’ competition with the former market leader for permanent monopoly profits. The ’neck to
Figure 24: The main mechanism in Liu et al. (2022).

Note: As interest rates get closer to the zero lower bound (ZLB), the value of a permanent monopoly position grows, spurring the market leader’s R&D investment. Close to the ZLB, the market follower understands that catching up on the market leader would put the two in a situation of fierce competition to capture the permanent profits. This moderates the increase the market follower’s R&D investment near the ZLB. As a result, the market leader invest much more in R&D near the ZLB and is much more likely to secure a permanent monopoly situation.
Acharya and Plantin (2019) argue that when interest rates are low enough, it is optimal for shareholders to use leverage to issue early cash payouts in the form of dividends and share buybacks (leveraged payouts). This happens when the cost of financing is less than the discounting applied by shareholders to future consumption. In this situation, the firm’s manager might act on behalf of the shareholders and bring future payouts forward using leverage. Higher leverage, however, would mean less skin in the game for shareholders, thus pushing them to exert less effort to ensure that the firms’ projects succeed, which would hurt aggregate productivity growth. Such a mechanism is especially powerful when the firm’s leverage is unconstrained, as would be the case when the firm does not rely on commercial banks for funding (capital markets, shadow banking). The authors argue that if such a mechanism is powerful enough, the optimal monetary policy may consist of leaning against the wind, i.e., not stimulating the economy, in order to contain leveraged payouts and maintain productive efficiency.
Kiyotaki et al. (2021) present a framework where persistently low interest rates can stifle investment, productivity and economic growth and even lead to a drop in the welfare of everyone in the domestic economy. Within this framework, a resource-constrained engineer/plant founder needs to sell the plant to finance investment. However, workers with human capital (engineers) cannot commit to providing the services required in the future to maintain and improve the plant’s productivity. As a result, engineers sell their services in a competitive labour market and receive their forward-looking marginal product, representing their contribution to future productivity. This reduces the new owner’s share of future revenues as more and more of the longer-term revenues are used to compensate engineers for maintaining productivity. However, the plant owner still has an obligation to pay fixed costs. This mismatch between long-term fixed costs and short-term revenue shares implies that a permanent drop in interest rates reduces the value of the plant from the owner’s perspective. Given that the value of the plant determines the financing available to fund productivity-enhancing investment, low-interest rates lead to a decrease in investment and productivity growth, thus reducing welfare (the model’s mechanism is summarised in Figure 25). The authors’ policy recommendation focuses on spurring productivity and economic growth by enhancing investment through investment subsidies. The investment subsidies are financed by taxing the engineers’ wages, as the sub-optimal aggregate outcome results from the engineers’ wages being inflated by the friction at the core of this model: the inability of engineers to commit future effort.

The mechanisms used in the paper above all focus on the firms’ behaviour in a low real interest rates context. Liu et al. (2022) differentiate between market leaders and market followers. On the other hand, the other two aforementioned papers adopt a representative firm approach and ignore supply sector heterogeneity. One, however, can imply what sort of firms the authors had in mind when building their models. For instance, Kiyotaki et al. (2021) focus on investment in new plants that might end up being credit constrained in a low-interest rates environment. On the other hand, Acharya and Plantin (2019) study the behaviour of larger corporate entities with access to the market financing required to finance externally shareholders’ payouts. What is the impact on aggregate productivity when larger, more established firms can better take advantage of low interest rates than their younger, smaller counterparts? According to Bai et al. (2018), the deregulation of the banking sector in the United States improved productivity outcomes by easing the financing conditions of more productive young firms, thereby fostering a more efficient allocation of production inputs. This finding aligns with the notion of a differentiated effect of low interest rates on the business population, which could potentially hinder the emergence of more productive smaller firms and consequently reduce aggregate productivity growth.
6 Managing Growth and the Keynesian growth synthesis

The canonical macroeconomic models of economic fluctuations do not have a role in explaining trends and patterns in productivity. Indeed, productivity trends are largely treated as exogenous. This means that the capacity of the economy is treated as a constraint on demand and, as a result, policy concentrates on ensuring that demand does not exhaust notions of fixed supply. This paradigm has dominated the approaches to monetary and fiscal policy in the post-war period. Indeed, attempts to boost activity by a succession of Chancellors, for example, by Barber, Lawson and Brown, have ended in busts rather than any sustained increases in prosperity. Yet, at the same time there is considerable evidence to suggest that aggregate productivity growth is neither fixed nor unaffected by well-designed policy interventions; that said, it is more of a question for the medium term rather than for the day to day of politics (Syverson, 2011).

There is a rich literature on endogenous growth, where the long-term growth of residual total factor productivity results from the economy’s agents’ (rational) behaviour.\textsuperscript{22} As indicated above, standard endogenous growth theory concerns itself with the long-term improvements in the economy’s capacity, as opposed to the study of economic management at business cycle frequencies through nominal interest rates. The recent macroeconomic practice has relied on what became known as the ‘new neoclassical synthesis’ in its attempts to bring aggregate demand in line with aggregate supply in a way to keep price inflation under control. The ‘new neoclassical synthesis’ combined the neoclassical intertemporal optimisation and rational expectations with the imperfect competition and the nominal rigidities of the New Keynesian models (Goodfriend and King, 1997; Goodfriend, 2004). Until recently, endogenous growth theory and New Keynesian economics represented two distinct traditions and interacted very little to study economic trends or help inform economic policy. However, things started to change after the global financial crisis after output in most advanced economies failed to revert to pre-2008 trends. As advanced economies emerged from the chaos created by severe financial disruptions, the L-shaped output recovery indicated that something has put these economies on a new trend for potential output and productivity. The standard framework used to manage business cycle fluctuations ignores the effect of shocks on long-term growth and found itself unable to answer the new set of questions brought by the post-2008 economic realities.

In the stylised representation of the standard framework outlined in Figure 26, the level of economic growth (\(\dot{g}\)) is independent of labour employed (labour force, labelled as \(l^f\)) or

\textsuperscript{22}See Aghion et al. (1998) and Aghion and Howitt (2008) for surveys of endogenous growth theory.
Figure 26: Exogenous supply with demand.

\[ \dot{g} \]

\[ AD1 \]

\[ AD2 \]

\[ G1 \]

\[ G2 \]

\[ \leftarrow \]

\[ \leftarrow \]

\[ \downarrow \]

\[ \downarrow \]

\[ f \]

Note: \( l^f \) is labour force and \( \dot{g} \) is the trend rate of economic growth.

other measures of economic inputs. However, aggregate demand will tend to increase with employment and may shift out with expansionary monetary and fiscal policies from AD1 to AD2, where AD signifies aggregate demand. In this setting, an exogenous fall in the economic growth rate from G1 to G2 would move the economy from A to B where there would be lower levels of factor employment. Unless, of course, if expansionary policies were used to boost aggregate demand and create a high employment low growth equilibrium at C. One may wish to interpret the period since the financial crisis through this lens.

Figure 27 further illustrates that if growth is also influenced by the demand for factor inputs such as labour, stemming from investment, trade, and public goods provision, then a downward shock to productivity growth can lead to an amplified reduction in both growth and labour employment. In this scenario, growth and labour employment may fall to point D. This is because growth is no longer exogenous to the factors employed. If, as in this case, the active labour force falls when demand falls it reduces trend growth, for example, from
the loss of firm-specific knowledge. But even here with appropriate demand management policy, the level of employment per se can ultimately be independent of the growth rate, as lower levels of economic growth can still provide full employment if aggregate demand rises to fully match employed labour supply – prior to Covid-19 and since the global financial crisis we had full employment with lower growth, and we look set to return to that level again. Even if the Covid-19 shock scars growth prospects then full employment can be obtained with expansive demand management policies and to an extent this is what we may have in prospect. But that high employment will not be accompanied by rapid increases in productivity and prosperity unless sustained policies shift up the endogenous growth path.

In this world, when there are demand shocks and an upward-sloping growth curve, output and employment volatility will tend to be higher. And, if we allow expectations of this variance to matter for the decision rules of firms, this may have then had an effect on ongoing investment as firms may defer their immediate investment plans, particularly if

23See Cerra et al. (2023) for a literature survey on the hysteresis effects on growth.
carrying debt in the face of uncertainty. Persistently poor productivity performance may well then be a low equilibrium trap. It is one that has arguably trapped the UK, particularly in some of its regions. The point is that a failure to control output fluctuations through active demand management may have permanent, or at least long-lived, consequences for growth and employment patterns. But if we consistently rely on aggregate demand to reach full employment, we may then run into other traps that we now see. These will hamper future demand management options, as fiscal space may evaporate, and monetary policy may get stuck at the Effective Lower Bound with quantitative easing acting to distort bond prices.

A new class of models are taking us away from a strict distinction between supply and demand. From the supply side, Guerrieri et al. (2022) and Baqaee and Farhi (2022) have explored a disaggregated model with multiple sectors, multiple factors, input-output linkages, downward nominal wage rigidities, credit constraints, and a zero lower bound. Here various complementarities can mean that a negative supply shock can be amplified through a further negative shift in demand. However, annoyingly, it does not follow, with this approach, that demand management is necessarily more effective. It depends. Fiscal policy needs to be targeted e.g., furlough or skills training and monetary policy directed at firm births and exits. This means a more granular approach to nurture a return to higher productivity. In Chadha et al. (2021), monetary and fiscal policy matter because they can combine to support financial conditions and support bank lending.

7 Government debt and fiscal policy

The global financial crisis heralded a sharp increase in the public debt-to-GDP ratios in advanced economies as current public expenditure rose sharply in 2010. The UK was particularly sensitive to the financial crisis as it focused on financial services, which accounted for some 10% of nominal GDP in 2010 (Chadha et al., 2017). Further increased public borrowing, following the Brexit referendum, during the Covid lockdowns and in the midst of the "cost of living crisis", which have combined higher spending without an increase in revenue performance means that public debt is likely to stay in the neighbourhood of 100% of GDP for some time. The objectives for low and stable public debt levels at under 40% of GDP now seem like a half-forgotten dream (OBR, 2023).

One possible reason for poor productivity performance is that higher current and expected levels of public debt may imply a fundamental and underlying deterioration in the fiscal position. In the long run that can only be closed by higher taxes on capital and and/or labour in order to provide a credible solution to the government’s present value budget constraint. If households and firms anticipate increases in taxes then labour and capital, as
factors of production, will expect to lose some of the incentive to be more productive because the post-tax return would fall. There is some form here. Cooley and Ohanian (1997), with careful calibration of a growth model, suggest that higher capital taxes in the UK during WW2 and in the postwar period played a significant role in relatively poor macroeconomic performance. They compare outcomes to a tax-smoothing alternative and find that such a strategy would have been significantly more preferable. The sharp increase in public debt as a share of income over the recent past may be leading to expectations of substantial increases in taxes.

To illustrate this point, the upper panel of Figure 28 shows the path of public debt against a five-year moving average of growth in income per head. The lower panel filters both series to remove high and low frequencies and suggests some correlation that might fit the hypothesis that public debt has played some role in acting in a deleterious manner on the growth in income per head, particularly as public investment has been regularly constrained with current expenditure dominating any fiscal impetus.

Public investment peaked recently at 3.3% of GDP in 2009 and has subsequently fallen to around 2.0% of GDP since 2010. Fiscal policy has been constrained by commitments to current expenditure and this has limited the room for manoeuvre under successive fiscal rules for greater levels of public investment. Switching fiscal policy towards more carefully assessed public investment to include infrastructure may make a great deal of sense, especially if the fiscal multiplier of public investment is greater than one as reported in many studies (Gechert, 2015).

8 Conclusion

This paper surveys several macroeconomic questions underpinning the underwhelming productivity performance of the United Kingdom. The paper raises questions rather than answers for the UK productivity debate, highlighting the need for a sustained research effort to understand the idiosyncrasies of productivity trends in the UK and possible ways to improve outcomes.

Modern growth accounting is usually the first step in analysing national productivity performance. This approach decomposes the performance of labour productivity into a capital deepening component and a (residual) total (multi) factor productivity component. Recent twists on this framework show that the post-2008 slowdown of the UK’s labour productivity can be fully accounted for by the anaemic growth of residual total factor productivity following the global financial crisis. Unfortunately, beyond the assumption that TFP is decided in some productivity frontier overseas, this does not allow us to say much about the UK’s
Figure 28: Debt to GDP ratio and productivity trends.

Note: The top panel represents the debt UK to GDP ratio and the 5-year centered averages growth of labour productivity in the UK between 1945 and 2019. The bottom panel represent HP-filtered growth series of debt to GDP and labour productivity in the UK (HP multiplier = 100). Source: Thomas and Dimsdale (2017), Office for Budget Responsibility (March 2024) and ONS (July 2022); authors’ calculations.
specific productivity performance. TFP is a residual where many factors are lumped together, including credit conditions, network effects, technology improvements, and firm-level organisational methods. Thus, allocating the UK’s poor labour productivity performance to weaker TFP growth keeps the UK’s productivity question wide open. Moreover, this framework cannot address what is potentially the central issue in the UK productivity debate: chronic underinvestment. We still need to establish the links between investment and TFP.

Public investment as a share of GDP dropped to a much lower long-term average after 1979, and the GDP share of business investment trended downward since the early 1990s. This is puzzling, given that interest rates have trended downward since the 1990s. The Solow-Swan growth model assumes that aggregate investment is an exogenous constant fraction of output and is, by construction, unable to relate capital investment to interest rate levels.

The Ramsey-Cass-Koopmans model derives aggregate investment from microeconomic foundations. In this model, the aggregate investment outcomes are linked to the level of interest rates and the growth of total factor productivity. We find that the RCK model’s household saving equation can match both the drops in interest rates and in total factor productivity after 2008 for certain values of the parameter driving the households’ elasticity of intertemporal substitution. The debates on the value of this elasticity in the literature are not settled, with microeconomic and macroeconomic studies pointing to different values. In any case, the experience from the post-2008 decade can inform us about how advanced economies adjusted to a loss in aggregate productivity growth that is likely to be repeated as these economies, including the UK, find their way toward greener modes of production and consumption.

The relative price of investment goods has been trending downward since the 1980s. This trend partly explains the decline of the UK’s investment share, as purchasing the same quantity of capital goods costs less in GDP terms. However, the UK’s underinvestment puzzle remains unsolved as UK business investment underperforms other countries with similar trends in investment goods’ prices.

Labour supply has increased, responding to a more flexible market favouring employers and increasing households’ indebtedness. Standard macroeconomic models lack a mechanism to reflect post-2008 trends in the labour market and therefore have missed the increase in aggregate hours. This increase in aggregate hours should have made capital more productive, spurring investment. However, this surge in investment never materialised, deepening the UK’s underinvestment puzzle.

Much of the recently revived ‘secular stagnation’ debate centred around the effects of population ageing in advanced economies. In the UK, population ageing was compounded by higher participation rates of older workers and higher unemployment rates among younger workers in the period between the global financial crisis and the onset of the Covid-19
pandemic. A more experienced workforce translates into higher human capital in national accounting. This and the increase in the share of workers with a university degree thus deepen the UK’s productivity puzzle.

Emerging literature provides theoretical arguments explaining the low investment despite low real interest rates. Liu et al. (2022) argue that low interest rates favour more industry concentration, which decreases competition and stifles investment in research and development. Kiyotaki et al (2021) argue that low interest rates might not help entrepreneurs invest in new projects because it increases the present value of their fixed costs more than the external valuation of the cashflow they may generate. This reduces the value of the new projects from the perspective of outside funders, thus reducing external financing and stifling investment. We therefore focus on the primacy of the decline in real rates and financial mechanisms for understanding low growth outcomes, which have over time built up debt relative to GDP as the demand for public services has grown. In this sense the lack of productivity growth has most likely fostered an increase in public and private indebtedness as a temporary solution in maintaining living standards and supporting public services.

A Derivation of growth accounting formulae

Both growth accounting formulae are derived from the definition of TFP as the output residual unexplained by changes in inputs

\[
\Delta \ln TFP = \Delta \ln Y - \alpha \Delta \ln K - (1 - \alpha) \Delta \ln N - (1 - \alpha) \Delta \ln Q. \tag{A.1}
\]

Note that on the right hand side of (A.1) can be rearranged to

\[
\Delta \ln TFP = \Delta \ln(Y/N) - \alpha \Delta \ln(K/N) - (1 - \alpha) \Delta \ln Q. \tag{A.2}
\]

Rearrange to get the standard labour productivity growth accounting

\[
\Delta \ln(Y/N) = \Delta \ln TFP + \alpha \Delta \ln(K/N) + (1 - \alpha) \Delta \ln Q. \tag{A.3}
\]

To obtain the growth accounting expression in Fernald et al. (2017), remark that the right hand side of (A.1) can be rearranged as follows

\[
\Delta \ln TFP = (1 - \alpha) \Delta \ln(Y/N) - \alpha \Delta \ln(K/Y) - (1 - \alpha) \Delta \ln Q. \tag{A.4}
\]
Divide by $1 - \alpha$ and rearrange to get the labour productivity growth accounting in Fernald et al. (2017)

$$\Delta \ln\left(\frac{Y}{N}\right) = \frac{1}{1 - \alpha} \Delta \ln TFP + \frac{\alpha}{1 - \alpha} \Delta \ln\left(\frac{K}{Y}\right) + \Delta \ln Q. \quad (A.5)$$
References


Oliver Wyman (2024). ‘Improving access to finance for small and medium-sized enterprises’, Levelling-Up Advisory Council, UK.


