Fixing the holes in economics: better theories for better growth

Working paper (May 2025)

Author
David Halpern CBE – President Emeritus, The Behavioural Insights Team, and Visiting Professor at King's College (London), El Erian Institute (Cambridge), and Imperial College (London).

president.emeritus@bi.team

This paper is inspired by discussions with Richard Thaler, Andy Haldane, Diane Coyle, Mohamed El-Erian, Nick Gruen, and colleagues in the Behavioural Insights Team and government. The author is particularly grateful for comments on the draft from Diane Coyle, Nida Broughton, Stephen Aldridge, Rachel Coyle, and Mark Byrne. The views expressed in this paper – and any errors – remain the author's own.

Contents

Summary

The core tenets of classical economics have been challenged over recent years. These challenges have been reflected in a string of Nobel Prizes for behavioural, institutional and methodological insights upgrading and overhauling classical models. The application of these insights, such as in pensions auto-enrollment, has proved highly effective. Yet, core economic policy as practised by many governments appears to have changed remarkably little.

This paper examines six areas where classical economic views have been overhauled over recent decades: (1) sentiment effects; (2) mental accounting and fungibility; (3) shrouding; (4) preference inconsistency; (5) 'irrational' cooperation; and (6) the role of institutions. The policy implications of updating the classical model are considered for each of these areas. A central theme is how information flows and is used. Government 'just getting out of the way' to let the market solve these informational issues often leads to non-optimum outcomes. The implications are extensive and point towards neglected policy levers with the potential to significantly boost economic growth and improve outcomes.

The world is currently going through a massive economic and geopolitical shock, itself in the wake of a quarter of a century of weak productivity growth with stagnant wages, pressure on public finances, and a struggle to deliver the quality of public services that citizens expect. Treasuries should seize the moment to explore the full range of possible responses, including testing new ideas to see if these can reignite a new phase of inclusive growth for their nations and people.

1. Introduction

> *Practical men who believe themselves to be quite exempt from any intellectual influence, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back.*
>
> John Maynard Keynes, *The General Theory of Employment, Interest and Money,* 1936

Over recent decades, there has been a string of Nobel Prizes awarded for ideas that have chipped away at core foundational elements of 'classical economics'. Many of these breakthroughs have been rooted in 'behavioural insights', improved econometrics, and a more evidence-based turn in economics: replacing the rational 'econ' with a more realistic model of the human to produce more accurate predictions and better theory. This includes the work of Nobel Laureates Kahneman (2002), Ostrom (2009), Shiller (2013), Thaler (2017), and Duflo et al (2019).

Despite this clutch of Nobels, most economic practice in treasuries and business departments across the world appears to have changed remarkably little. One explanation is that 'practical men' (and women) developing economic policy don't really draw much on economic theory, but instead draw on some other pool of knowledge or experience, including empirical but relatively atheoretical work.[1] A more worrying alternative is that such practitioners, be they Politicians or officials, are too often still 'slaves of some defunct economist', and a major reason for the disappointing results of much economic policy is that it is based on ill-founded premises.

This essay brings together six broad examples of where economic theory has been substantially challenged or updated, but (arguably) policy has yet to be correspondingly updated. For each, there is a brief summary of the evidence, and policy implications suggested. This paper draws heavily on the UK context, building on previous work, but might also be applied to other nations with underperforming economic policy, such as Japan and Italy.

While the primary focus of this paper is on 'market failures', there is a corollary thread of 'government failures'. Indeed, the real concern is the overlap between the two, or what we might call government accounts of market failures.[2]

2008 and the aftershock…

> *In the run-up to the crisis, the Federal Reserve Board's sophisticated forecasting system did not foresee the major risks to the global economy. Nor did the model developed by the International Monetary Fund, which concluded as late as the spring of 2007 that "global economic risks [had] declined*".
>
> Alan Greenspan, *Never Saw It Coming: Why the Financial Crisis Took Economists by Surprise*. Council on Foreign Relations, 2013.

In the wake of the financial crisis of 2008, there was much soul-searching in economic policy institutions, including the IMF, central banks, and governments. Not only was there a major recession, but it was widely seen to have been rooted in a failure of economic policy and prediction.

These soul-searching conversations were certainly occurring in the UK, not least led by the then Cabinet Secretary (2005-2011) Gus O'Donnell. He was particularly well qualified to pick up on the issue, not just as Britain's most senior civil servant, but as a former head of the Treasury, with previous senior roles in the IMF, World Bank, as an academic economist at Warwick and Oxford.

O'Donnell was clear that the models had failed to predict the crisis and needed a major overhaul. This led him to be a pivotal supporter behind the creation of the UK's Behavioural Insights Team (or 'Nudge Unit'). Nonetheless, he later reflected that despite the crisis, the Treasury seemed to go back to 'business as usual'.

It was an observation not unique to the UK. After a couple of years of hand-wringing, most of the big economic policy institutions appeared to go back to models and policies that were substantially unchanged. There were some policy changes, such as some enhanced capital requirements for banks, intended to reduce the likelihood of future contagion and government bail-outs. Yet the profound failure and evident falsification of core economic assumptions did not seem to lead to a wider overhaul.

Anomalies

The financial crisis did not stand alone. Economists such as Ostrom, Shiller and Thaler had been documenting severe violations of the 'classical' economic model for at least the previous two decades. Thaler assembled a series of these 'anomalies' in a long-running series of papers in the Journal of Economic Perspectives, later brought together in an edited book, *The Winner's Curse: Paradoxes and Anomalies of Economic Life* (1992).[3]

Though Shiller and Thaler gradually rose in prominence, much mainstream economic teaching and practice remained stubbornly impervious to the evidence they were assembling. The conventional defence was generally that these anomalies were either minor, or that even if they were more widespread, they would broadly average out over time and markets. In other words, even if many or most individuals aren't rational utility maximisers, on average they would act as if they were – an emergent blind hand of the market producing an econ after all.[4]

The rise of a challenger theory out of a series of anomalies is what philosopher of science Kuhn would have recognised as setting the conditions for a 'paradigm shift':

> *Some revolutions are large, like those associated with the names of Copernicus, Newton, or Darwin, but most are much smaller, like the discovery of oxygen or the planet Uranus. The usual prelude to changes of this sort is, I believed, the awareness of anomaly, of an occurrence or set of occurrences that does not fit existing ways of ordering phenomena. The changes that result therefore require 'putting on a different kind of thinking-cap', one that renders the anomalous lawlike but that, in the process, also transforms the order exhibited by some other phenomena, previously unproblematic.*
>
> Thomas S Kuhn, *The Essential Tension* (1977).

Indeed, Thaler quoted Kuhn in his introduction to The Winner's Curse, that 'discovery commences with the awareness of the anomaly'. Thaler went on to say that he hoped the book would accomplish that awareness, and that it would lead to 'the development of the new, improved version of economic theory'.

Thaler did go on and get the Nobel, albeit 25 years later. He mused publicly and privately that 'the lunatics are running the asylum'. Behavioural economics can

certainly claim to have achieved mainstream recognition. It is now increasingly widely taught, if not universally accepted.

At least some of this has found its way into policy. The most prominent example is the adoption of auto-enrollment in pensions, an idea which Thaler himself worked on and championed from the mid-90s. Viewed through a classical lens, it shouldn't matter whether a person has to opt in or opt out of a pension. If the person will be better off in the pension scheme – because the employer and government will substantially top-up their saving – then they will join the scheme (unless for example, they have a rare medical condition that they know will lead to an early death, or perhaps a heavy smoking habit that they have no plans to quit…) Yet we now know conclusively that switching from an opt-in to opt-out system is a game-changer in terms of increasing savings (see box).

However, across the wider sweep of economic policy, the impact of this so-called Kuhnian paradigm shift in economic theory is surprisingly hard to detect. In this sense, it is more akin to the account of Lakatos, a rival to Kuhn. For Lakatos, experts are able to resist widespread falsifications or anomalies. They are able to do this because any theory or observation is composed of a large bundle of assumptions and beliefs. So when an apparent falsification comes along, like one of Thaler or Shiller's anomalies, existing theorists can always blame it on a 'bridge assumption' while protecting the core elements of their theory. For example, an economic policymaker might argue that the pensions' auto-enrollment result is a result of a very narrow circumstance, and has little bearing on economic behaviour more widely. As such, the elegant simplicity of classical economics can remain the guiding principle of most economic policy.

Pensions auto-enrollment: a game changer that doesn't fit with the classical economic model

On both sides of the Atlantic, governments have spent billions of dollars through tax subsidies to encourage workers to save more into their pensions. However, despite the generosity of these subsidies, the majority of workers did not take up work-based pension schemes.

An influential analysis by Chetty *et al* of the Danish pension system estimated that every $1 of tax subsidy led to a miserly $0.01 of extra saving. Around 85% of savers seemed unaffected by the subsidy, while around 15% responded by aligning their savings to be more tax efficient, but not saving any extra.

Building on the work of Thaler and Benartzi, a number of countries shifted from an 'opt-in' to an opt-out model of pension saving, including New Zealand, the UK and (in part) the US. In contrast to the ineffectiveness of tax subsidies, changing the default was spectacularly successful. Around 91% of eligible workers stuck with the default to save more, bringing in tens of millions of new savers, including younger and lower-paid workers.

Curiously, governments have generally not reduced the large and ineffective tax subsidies for pension saving that run to tens of billions a year. Nor have they fully brought in the other aspect of the original work, which is to have an 'auto-escalation' mechanism whereby, when workers get a pay rise, a slice of that rise is used to increase the percentage of savings going into the pension.

Economic anomalies and their policy implications

> *I have come to believe that people, especially during periods of extreme economic stress, act in ways that are more predictable than economists have traditionally understood. More important, such behavior can be measured and should be made an integral part of economic forecasting and economic policy making.*
>
> Alan Greenspan, *Never Saw It Coming: Why the Financial Crisis Took Economists by Surprise*. Council on Foreign Relations, 2013.

Though Greenspan may have been as wrong-footed as many others by the financial crisis of 2008, and the failures of economic models to predict or prevent the crisis, he at least deserves credit for updating his own views.

In this, the main section of this essay, we will explore a series of contemporary anomalies and policy failures. The point is not ridicule. Treasuries and government Departments that deal with economic policy are full of talented, smart, and dedicated staff. A failure to update their models is not because they are stupid or overtly ideological. Politicians may seem easier targets, with their more

ideology-based approaches to policies. But they, too, are generally smart people with good intentions to boost growth. What politician wouldn't want to improve economic outcomes?

Yet good intentions aren't enough. Humans, even the very smartest of us, are prone to overconfidence, and to an 'illusion of explanatory depth'. Issues, questions and practices that feel familiar to us bring a halo of being understood, even though when quizzed, people often are shocked to discover that their accounts are surprisingly thin or implausible. Too often, we navigate the world confidently, but wrong – our minds skilfully bridging the gap with rationalisations. The smarter a person is, the better they can be at finding one of Lakatos's 'bridge assumptions' to take the fall.

2. Sentiment – talking up (or down) markets

> *We need to incorporate the contagion of narratives into economic theory. Otherwise, we remain blind to a very real, very palpable, very important mechanism for economic change, as well as a crucial element for economic forecasting. If we do not understand the epidemics of popular narratives, we do not fully understand changes in the economy and in economic behaviour.*
>
> Robert J Shiller, *Narrative Economics: How Stories Go Viral and Drive Major Economic Events*

The UK government would appear to have had more than one recent, and bruising, lesson on the importance of 'sentiment' in markets, or what Keynes famously called 'animal spirits'. Among the clutch of disruptive economists and Nobels, it is Bob Shiller who has done the most in recent years to highlight the importance of these 'sentiments' or 'narratives' to economies.

The basic idea is simple enough. Economic behaviour is strongly affected not only by the private assessment of the economic options by individuals, but also by their assessment of that of others. Is it worth investing half your salary in buying tulip bulbs or bitcoins? It might be, if you think everyone else in the world thinks so, and especially if you think you can get out before the crash. In more normal times, it can also be a good strategy, piggybacking on the knowledge of others, not least expressed through the growing popularity and decent performance of passive funds (themselves partly inspired by the work of Nobel winner Fama and the efficient market hypothesis, see below).

Market sentiment is an important and rational consideration for businesses, too. The financial well-being and likely prospects of customers are extremely relevant to a business trying to decide what levels of stock to order (for example, before Christmas) or whether to invest in additional capacity to increase production.

The problem is that such sentiment or 'animal spirits' can lead to massive overswings in the form of bubbles and busts. Sentiment can be self-fulfilling, at least for a while. If others around you think things are going to get better – or a price will rise – they will

also buy stock, invest in a plant, and take on extra people, all of which will boost incomes and activity. Indeed, at the time of writing, markets have shown massive volatility, partly driven by the actions of Trump, but further amplified by rumours and sentiment-based swings.

Whatever the 'fundamentals', betting against market sentiment can be a dangerous game. But when this positivity becomes disconnected from reality, such as the infamous 'tulip mania', then it becomes an unsustainable bubble. When this bubble busts, many people are ruined, and economies can be scarred for a generation. As was noted in the wake of the Great Depression:

> *Everybody's opinion is largely guided by the opinion of everybody else, even the people with the coolest heads will at least "fear the fears of other men" and contribute to the panic of which such fears are a part.*
>
> Irving Fisher, 1932

Note that it is perfectly 'rational' to factor in the sentiment and financial prospects of others in the economy. The trader who correctly predicts that the economy is rising and buys extra stock while supplies are plentiful and cheap will prosper. Similarly, the overconfident trader who buys too much stock before the market crashes will suffer. The issue is complicated further by systematic biases in our ability to forecast, particularly by relatively naive investors (though not limited to).

Evidence

There is now extensive literature on the impact of sentiment on economic growth. Many of these studies indicate substantial effects.

Melani (2017) concluded that variations in sentiment are responsible for more than 40 percent of US business cycle fluctuations. Flynn and Sastry (2024) concluded that 'narratives explain 32% and 18% of the output reductions over the early 2000s recession and Great Recession, respectively, and 19% of output variance'.

One well-studied example is how the qualitative sentiment in the language of central bankers independently predicts market movements over and above formal changes in rates and other measures. These include impacts on inflation and industrial production. Similarly with the USA, tonality in Fed forecasts predicts growth over and above formal bank predictions and rate changes up to 4 quarters ahead

(Sharp et al, 2023). Recent work has even shown how body language and facial expressions of policymakers affect markets, such as when central bankers are talking about monetary policy when giving evidence.

Over and above the headline results, there have been attempts to pull apart these effects in more detail. For example, sentiment effects on stock market returns have been found to be especially pronounced when firms themselves are more uncertain (ie sentiment plugs gaps left by more confident predictions or uncertain times). Some analyses have suggested that sentiment effects appear to be particularly pronounced outside G7 countries, where sentiments move not only prices but also underlying productivity.

Another important detail is that sentiment effects could be asymmetric. Sharp et al (2020) concluded that 'much of tonality's [sentiment expressed in Federal Reserve forecasts] forecasting power arises from its signal of downside risks to economic performance and stock returns'. These tonality effects are also more impactful during times of uncertainty. One interpretation of this may be that policymakers are generally wary about talking down an economy – companies, governments and Ministers are generally in the business of talking up the positive impacts of their policies and countries. So when they do express reservations or concerns, these are more striking, unexpected and impactful.

One recent meta-analysis by Gric et al explored whether the literature on sentiment effects might be distorted by publication bias – the disproportionate publication of dramatic and large positive effects relative to null results. (Appreciate the irony: evidence on sentiments affected by sentiments…) The authors did indeed find some evidence for publication bias. Controlling for this bias did reduce average effect sizes, but sentiment effects were still found to be robust. This meta-analysis also concluded that the sentiment of individual investors matters more than that of institutions, and interestingly, is more pronounced in US markets than in Europe.

Implications

Central bankers and even large firms have become increasingly aware of sentiment effects. There is a growing awareness that how and what they say matters as well as what they 'do'.

This has led bankers and business leaders to be increasingly careful in the wording they use. Carney's doctrine of 'forward guidance' can be seen as a living example of this, with an overt intention to inform and steer market sentiment.

Within the business world, AI has been added to the arsenal of market analysts to study the tone and sentiment and wording of statements by firms and bankers. In response, many of these institutions have become more careful in their worded statements. This, in turn, has led some sophisticated market analysts to shift their focus to less structured comments and responses, such as CEOs' responses to questions at AGMs in the hunt for predictive advantages.

One very practical implication for central bankers, ministers and CEOs is to systematically test how people will interpret and respond to the guidance and evidence they put out before they put it out, and they adjust it accordingly. For example, BIT ran a series of randomised control trials testing variations in the format of inflation guidance put out by the Bank of England. Significant differences were found in comprehension and sentiment. For example, people were more likely to correctly understand visual graphics and concrete examples than general statistics. Interestingly, this increased comprehension was in turn associated with higher trust in the bank as an institution (noting the importance of trust itself).

Presumably, the intent of a Central Bank is to improve the calibration, or empirical grounding, of people's expectations of what the economy will do. This should help to avoid bubbles and busts grounded in 'frothy' sentiments, as well as reassuring business and consumers that the Bank will act to keep the parameters of the economy within certain guide rails. This, in turn, should strengthen some useful self-fulfilling prophecies, such as encouraging businesses and workers to show wage restraint, confident in the knowledge that runaway inflation will be prevented.

Some policymakers and business leaders may wish to go further. A President, Prime Minister or Chancellor who can convince business leaders that their growth plan will work, or that the winds of change are in the right direction, can leverage this sentiment to boost additional investment and growth. Indeed, Trump and some market players in the US appear to be overtly pursuing this strategy.

Of course, the best boosts to sentiment may come from authentic plans and getting the 'fundamentals' right. Governments that bring forward convincing plans for how they will cut their deficits, including spending restraint, may reasonably be expected

to be rewarded with lower bond yields (cheaper borrowing), which in turn may ultimately lessen their need to deliver such cuts.

Within the UK, Prime Minister Truss's hubris clearly failed to deliver this sentiment boost around her disastrous 'mini-budget', not least as she sought to bypass the judgment of the independent OBR. On the other hand, the 2024 Labour government appeared to be rather too good at talking down the markets. Whether Trump will succeed is yet to be seen – he has certainly benefited short-term traders who correctly anticipated the oscillations in sentiment he induced.

At the very least, senior parts of governments need to learn the lesson that sentiment matters greatly. Crass though it might seem, they would be wise to systematically test and revise the wording and plans they release to be attuned to sentiment, noting that this testing can also extend to whether thought leaders are persuaded by the underlying proposals and evidence.

Second, it reinforces the case for building up trust and confidence in economic institutions as a good in itself. Central Banks are right to worry about whether businesses and people trust their judgments and actions. Similarly, governments are right to consider what institutions may be missing to boost confidence (as well as competence), including the independence of those institutions.

Finally, it is worth considering, and potentially refining, the basis and transmission of so-called business and consumer confidence surveys. We know people are prone to systematic error in predictions, including through the influence and estimates of the sentiment of others. But we also have increasing knowledge about how people can improve their 'calibration' or prediction. These understandings could be used to produce more robust and better calibrated 'confidence surveys', which would in turn become more predictive, valued, and would act as stabilisers of 'grounded-growth'.

3. Mental accounting – imperfect fungibility

> *The problem seems to be that while economists have gotten increasingly sophisticated and clever, consumers have remained decidedly human. This leaves open the question of whose behaviour we are trying to model.*
>
> Richard Thaler, The Winner's Curse, (1992. p121).

In late 2024, Richard Thaler and I led a seminar in Downing Street chaired by the Prime Minister's economic policy adviser, Rav Athal. Richard and I had talked about the session in advance and decided that a good focus would be on areas where policy had failed to catch up with current economic theory and evidence. Top of that list, for Richard, was mental accounting.

For readers not familiar with the concept, here is a short summary. In classical economic theory, a dollar is a dollar, and a pound is a pound. It doesn't matter who gave it to you or where you store it: its value is the same. Rational 'econs' understand this. Humans, on the other hand, seem to have a different view. Humans often put money into mental, literal or digital 'jam-jars'. For example, one jar is for rent, another is for the kids' education, and another is for fun nights out and other treats.[5]

The problem with jam-jarring – or mental accounting – is that it messes up a basic principle of economics that a pound is a pound. For example, imagine if a person's 'fun' jar is running low, but they've just heard that their best friend, whom they haven't seen for ages, is in town. There's plenty of money in the rent jar, and it's not due for a couple of weeks, and their payday is coming up. To an econ, this isn't much of a puzzle. Use some of the money from the 'rent jar' to see your friend (a high pleasure return per pound), since you will have enough money for your rent after you get paid. Indeed, for the econ, why have these 'jars' at all?

For humans, this is more difficult. 'But that's the rent money!' As humans ourselves, we recognise where the concern might come from. We worry that we might not have the self-control to put aside that money when it comes in. Maybe next week another friend will come into town. Or there will be a great movie on. Humans use

'jam-jars', or mental accounting, as part of our everyday strategies to control temptation, and simply organise our lives and our finances.

Thaler included mental accounting, or 'fungibility', as one of the anomalies that he highlighted in The Winner's Curse. He noted overwhelming evidence that people appeared to put money into separate mental accounts for savings, housing, and day-to-day expenditure. This led to 'irrational' behaviour insofar as people could often do financially better by rearranging money across these pots. A common everyday example of this is where people borrow money at high interest rates, while having money in bank accounts that is earning much lower interest rates. Many businesses do this too, such as setting aside budgets for training or IT for a year or more, and that may not get revisited even if the opportunity costs shift.

Evidence

An immediate question is how mental accounting might affect policy-relevant behaviours across different aspects of financial life. For example, one of the arguments against auto-enrollment for pensions is that, even if it increased pension savings (by reducing friction), overall it wouldn't make much difference because if people put more money into their pension, they would simply reduce their savings somewhere else. Hence, people might make smaller capital payments on their mortgages, or reduce their savings in interest-earning accounts or equities.[6]

However, the evidence is that this is not the case. Studies in both the US and UK have concluded that, even among low earners, auto-enrolment did not displace other savings (though there was UK evidence for a small short-term effect on increased debt defaults among low earners as people adjusted, alongside an increase in taking mortgages, suggesting pension saving may have prompted people to make other financial decisions). Indeed, this lack of substitution is sometimes used as a critique of auto-enrollment – that some low earners are defaulted into putting aside money for pensions when they would do better to use the money to reduce high-interest debts. In this case, one violation of rational economic behaviour (the power of defaults) could be argued to cancel out another (mental accounting), such that policymakers would have been better off leaving these people alone.

A more sophisticated approach would seek to account for the range of constraints and biases. In the case of those with low and volatile incomes, it might mean that

policy should respect and echo people's mental accounting by defaulting saving not just to pensions, but to linked 'rainy day' saving accounts (which they in turn could reallocate between if they wished).

Viewed through a classical lens, the addition of 'rainy day' savings doesn't make much sense. It is either reducing the financial flexibility (by putting money into a pot) or, if the friction to access the money is low, then what's the point (unless it is getting a high and attractive interest rate – then why not just offer that)?

When viewed through a human lens, rainy day savings accounts potentially offer large benefits. These benefits rest precisely on why people find mental accounting helpful. In this case, having some money in an account for emergencies serves as a 'commitment device' to ensure it is there for when their car breaks down or the washing machine needs replacing. These benefits are even larger once so-called 'scarcity' effects are layered in: that when people are under financial stress, they find it harder to make good short-term financial decisions and, rather importantly, having a rainy day saving pot frees up mental capacity that was otherwise taken up with worrying about the next household financial shock.

Winter fuel payments (WFP)

One of the early and unexpected political shocks to the 2024 Labour Government was around reactions to an announcement to cut the UK's 'winter fuel payment' for older people.

Within Whitehall, the Winter Fuel payment has long been seen as a possible target for cuts. It was a sizable £1.5 billion, and was relatively poorly targeted since it was also paid to affluent pensioners. Within BIT, we examined options for a previous Government, including making it easy for wealthy pensioners to opt out or donate the allowance and enabling it to be amassed into home improvements to lower fuel bills. Furthermore, pensioner poverty had substantially reduced since the Winter Fuel Payment was introduced and – if it were to be kept – it could be amalgamated into other benefits, including ongoing rises in state pensions. Finally, there were other more targeted options that could be deployed, such as the 'Warm Fuel Discount', a needs-based discount on fuel bills.

So when the Chancellor felt under pressure to give examples of the kind of 'tough choices' she was prepared to make to improve public finances, the Winter Fuel

Payment was a seemingly obvious target. The cut was announced in late July 2024. Yet the Chancellor and the Treasury seemed genuinely surprised at the strength of the backlash against the cut. A substantial part of this backlash can be linked to the power of mental accounting.

As it happens, there was compelling pre-existing evidence of mental accounting linked to the Winter Fuel Payment. Eligibility for the WFP depends on age (having someone in the household over 60). An IFS study (2011) used the UK's household expenditure survey to compare differences in spending on fuel between those just above and just below this eligibility threshold. Given that only around 3% of such household expenditure was spent on fuel at that time, and there was no conditionality attached to the payment, from a classical viewpoint, this is roughly the proportion that we might expect a household getting a windfall extra £300 might spend on fuel. In fact, the study found it was more than 10-fold this level: 47% of the WFP was spent on increased fuel consumption.

At the same time, the cut was announced separately from planned increases to the state pension. Even though these announcements followed not long after (in the budget and before winter), people still experienced money being taken from the winter fuel 'jar'. It should not have mattered that, from a classical point of view. But with humans using mental accounting, it was disproportionately painful, including politically.

The apprenticeship levy

It is not only individuals who appear to engage in mental accounting. Firms appear to do much the same, or at least the people who run them.

The apprenticeship levy illustrates the issue. The levy requires firms with payrolls of greater than £3 million pa to contribute 0.5% to the levy. However, they can offset up to £15,000 of expenditure on training against the levy, and specifically the support of apprenticeships. Note that for a company with a payroll of £3 million, 0.5% is exactly £15,000.

Many in the Treasury were sceptical about the design of the levy. However, they later concluded that it was 'unexpectedly' effective. It had long been noted that UK firms, in comparison to those countries such as Germany and Austria, spent far less money or effort in training up workers in the skills that they needed. This was often

attributed to a combination of (rational) concerns about 'free-riding' – that firms might steal trainees from others rather than paying for it themselves, together with a 'cultural' disinterest in training and development.

In behavioural terms, whether intentionally or otherwise, the levy encouraged the creation of a mental (and literal) account for apprenticeship training. Once crystallised as a levy, the £15,000 became seen as that firm's own money in a 'jar' for training. This harnessed mental accounting, and an allied 'endowment effect' (it is roughly twice as painful to lose a pound or dollar of value than to gain it).

An alternative to the levy would have been an unlabeled payroll tax, or simply any tax, alongside incentives or payments for training. But it is unlikely this would have had the same impact on the thinking and behaviour of firms.

The apprenticeship levy is not unique. For example, R&D tax credits have a similar broad structure. Firms can claim back tax credits for R&D activity, with the intention of increasing this activity (itself partly based on evidence from the US and elsewhere of efficacy). A classical account justifies such credits as addressing the spill-over effects of R&D activity, due to knowledge spill-overs that can drive specialist skills, innovations and raise firm performance more widely (even if firms intend to capture the gains through patents and confidentiality agreements).

Yet R&D tax credits almost certainly also operate through reinforcing mental accounting, creating a 'jar' or percentage target that firms become more aware of with respect to R&D activity. That said, R&D credits are somewhat let down by elements of their administration that undermine this effect, such as ambiguities in what activities will be accepted as R&D, and delays and accounting rules that often lead to the R&D tax rebate being put into a different account.

Implications

There is overwhelming evidence that people think about money in 'mental accounts'. This is an important violation of the classical model and of the presumption that money is fungible.

Rather than wish this away as an annoying distortion, Treasury and government policymakers should map and understand these mental accounts. In this way, pounds and dollars can be applied with greater marginal impact, even if in classical

terms, it should be for the economic agent to figure out this marginality for themselves.

Policymakers wishing to promote growth might wish to make more overt the particular mental accounts, or 'jars', that are disproportionately important for a nation's long-term growth. We partly do this through the language of pensions and investments. But we might wish to do this more overtly. For example, Governments might publish alongside budgets the implications for current versus future spending for different cohorts, or more overtly separate 'fun' or 'sin' taxes from 'essentials' and 'investments' for the future.

Some of the nations that have achieved the most successful growth in the post-WW2 period did so with a combination of great collective effort and a tacit embrace of mental accounting. For example, when Asian tigers built their economies – South Korea in shipbuilding, Japan in car manufacturing and electronics, or Singapore in trade and refining – their industrial strategies were matched with collective restraint in consumption and a concentration in investment. Replicating and leveraging this type of mental accounting in individuals, firms, and institutions may appear messy from a classical view, but it may be more realistic and effective in terms of the grain of human cognition.

4. Shrouding – if only markets were perfectly efficient

> *A shrouded attribute is a product attribute that is hidden by a firm, even though the attribute could be nearly costlessly revealed.*
>
> Gabaix and Laibson, 2005

In a world of perfect markets, there would be no transaction costs, consumers would be able to immediately distinguish between product qualities and values, better companies would rapidly grow, and excess profits would be choked off.

The efficient market hypothesis (EMH) is a more specific and beautiful idea. It states that asset and other prices accurately reflect available information, for which Fama et al won the Nobel in 2013. It predicts that new information affects prices almost immediately, such that it is very difficult to 'beat the market'. It promises efficiency, growth and fairness (of a sort) because it implies a world in which good products, services and ideas displace the less good and inefficient.

In a world of efficient markets, the invention of a cheaper form of lighting (cf the electric lightbulb) or the development of a breakthrough drug is rapidly recognised by consumers and investors alike. Capital and sales flood into the new product or process, and away from its less effective rivals. For example, as news circulated of Edison's electric light bulb, the value of his company leapt, while the share prices of gas companies tumbled. Edison became rich, and householders saved money.

Technological innovation, twinned with efficient markets, can deliver extraordinary productivity increases. Edison's light bulbs were around seven times more efficient than gas lighting, just as gas lighting was much more efficient than the oil lighting it displaced. This twinning of markets and technology helped drive the 30,000 times increase in efficiency with which we now produce a lumen of light compared with the year 1800 (as measured by how many hours of labour are required).[7]

It is sometimes stated that efficient markets require everyone to be a rational economic maximiser, with perfect information. Yet that is not right – the model is more subtle and 'efficient' than that.

Harvard economist David Laibson has illustrated this market efficacy with the example of a naive and hungry visitor to the US. Even if the visitor had no knowledge of prices or the value of the dollar, they would still be very unlikely to be offered a banana for $1,000. Though the unwitting visitor might have paid this price, the efficient market ensures that there aren't vendors waiting at JFK airport to sell such outrageously priced fruit. The EMH just needs a reasonable number of consumers to shop around to establish prices, and the outrageously expensive fruit store will either have to drop its prices or go out of business.

Unfortunately, even in highly competitive markets, a version of the overpriced fruit arises frequently. An everyday example, Laibson cites, is how banks have for decades made a lot of money from 'hidden charges', such as on overdrafts. The puzzle for the EMH is to explain why challenger banks, offering transparent, good value charging, fail to displace such 'bad' banks and wipe out their bad practices. Laibson highlights how, even in highly efficient markets:

> *… two kinds of exploitation coexist. Optimising firms exploit naïve consumers through marketing schemes that shroud negative product information. In turn, sophisticated consumers exploit these marketing schemes. It is not profitable to try and lure either of them to non-exploitative firms. As a result, the distortions due to consumer biases persist across a wide range of markets.*
>
> Gabaix and Laibson, 2004

In more everyday language, there are plenty of naive or myopic customers which is most of us when we are in a hurry (do you read the small print?). We also mispredict our own behaviour: 'I'll claim that cashback' or 'I'll plan ahead and will avoid getting that overdraft'. The challenger bank that is open about charges and offers a simple, transparent fee structure will fail to attract these myopic customers, as its product will look more expensive. Instead, they will be drawn to the cheaper headline prices and teaser products of the 'bad' bank.

What about the sophisticated consumers? Surely they would choose the transparent challenger bank? No. If they have correctly understood the hidden charges, they can avoid the overdraft and get a banking product at a discount rate. In effect, they have access to a product that is being subsidised for them by their more myopic neighbours.

The net result is that both naive and sophisticated consumers stick with the bad bank, and the transparent challenger fails.

Evidence

First, let us recognise that there is good evidence for the benefits of markets that allow prices to adjust according to information. For example, there have been massive increases in the efficiency of the steel industry, with an 80 percent drop in the energy required to make a ton of steel from 1900 to 2010. Within market economies such as the US, technological innovations that made steel cheaper were rapidly adopted as prices moved and capital shifted. However, in non-market economies, this process was inhibited, helping to explain 20-year lags in the adoption of such technologies.

The Efficient Market Hypothesis itself helped inspire the massive growth of passive funds, which now run to trillions of dollars and own large proportions of global stocks – backed by the increasingly robust killer finding that passive funds outperform the vast majority of active managers (helped across the line by the lower fees of passive funds too).

Despite this, the type of market failures identified by Laibson et al appears to be widespread. The invisible hand of the market, populated with a mix of humans and a few 'econs', regularly finds a stable equilibrium that regularly exploits and conceals. Viewed from the side of the hard-working entrepreneur, the better value or quality service will often fail to break through.

These effects can be seen in periodic studies of 'consumer detriment', where regulators seek to identify excess profits and overcharging across sectors. Classic examples tend to be in sectors with strong informational asymmetries (eg, car repairs – do you know if your filter needs changing?) or those with distressed consumers (such as funeral parlours and veterinary services).

The strength of these effects is illustrated through a series of BIT studies. For example, we asked people to choose between a selection of deals offering foreign exchange, and incentivised them to find the best deal. These adverts varied in the degree of transparency. Sure enough, when presented with a set of deals with more transparent pricing, this roughly doubled the number of people able to identify the best option. But when a mixture of transparent and more opaque options was

offered, akin to a realistic market, performance collapsed. The opaque pricers profited nicely.

Similarly, prospective consumers were exposed to a series of adverts by fictional companies to test the effect of 'greenwashing'. We found that adding in images of a tree or greenery made consumers much more positively inclined towards those companies, even without any substance behind the image. Strikingly, this greenwashing was especially effective with consumers who cared more about the environment.

These effects play out in real markets, too. In a classic study of plumbers, those who called themselves 'AAA plumbers' and other similar names to get to the top of old-school phone books, or who listed themselves under several different names, were found to charge more and deliver worse service. You can see why it works. If the sink was leaking and you were in a hurry, you wouldn't go through an elaborate search to find a plumber. In this and other cases, the actual market (versus the 'efficient' market) means that these rogue businesses flourish.

There is a strong case that shrouding problems contribute to the long-run fall in 'allocative efficiency' across market economies, which explains between a third and a half of the drop-off in productivity and economic growth over recent decades. For example, they likely contribute to the slow diffusion and adoption of new technologies and to the reduced churn in the labour market, including that most job moves, when they do occur, are ones at similar productivity levels.

We are not arguing here that shrouding is the only factor behind this fall. Economic policies on both sides of the Atlantic have protected incumbent but inefficient small and large firms from failure. In particular, the extended period of low interest rates led to a rise in 'zombie firms' from 4% in the late 1980s to 15% by 2017 across OECD countries. At the same time, the rise of 'superstar' firms with winner-take-all characteristics, high mark-ups, low labour demands, and that use their market power to hold out challengers has further lowered productivity growth rates. Though exacerbated by the 2008 financial crisis and Covid-19, as governments and central banks lowered interest rates and sought to protect firms, this blunting of market competition appears to have preceded the 2008 shock, at least in the US.

One response is to argue that a sustained rise in interest rates, alongside antitrust measures, should lead to a cull of weaker firms and recovery of productivity growth,

albeit at a likely cost of higher unemployment. However, cranking up interest rates to kill off weaker firms is akin to brutal chemotherapy to heal the body, on the basis that the toxin will hopefully differentially kill off the bad cells faster than the good ones. But a better or more targeted approach would be to improve market functioning and informational flows by increasing transparency and reducing shrouding, so that more productive firms would grow more rapidly.

The introduction of tariffs and the current trade war are likely to worsen these problems, as well as drive up costs. Companies that export are substantially more productive. At the same time, across Europe and the UK, market competition is beset by shrouding problems and often by efforts to protect small and inefficient companies.

Finally, these shrouding effects are made worse by failures in a particular market that ought to help: the advice market. Almost by definition, people will tend to seek advice when they are unsure how to decide, and a strong information asymmetry will exist. How then is the consumer to distinguish good from bad advice? Furthermore, if good advice is more expensive or difficult to deliver, there is a high risk of adverse selection, where poor quality (potentially 'free') advice crowds out better quality advice. This adverse selection is made even more likely if people overrate their understanding or abilities, and especially if those of lower objective ability or understanding overrate their ability (as shown in the Dunning-Kruger effect).

The key role of market intermediaries

Advice products and services do exist: switching sites (eg, for insurance, energy); financial advisers; Consumer Reports (US) and Which? (UK); and review platforms. However, most consumers appear reluctant to pay for advice (eg less than 0.5% pay for access to Consumer Reports or Which?), and crucially face shrouding issues around advice (eg, fake reviews).

In some sectors, there are now well-established market intermediaries with sustainable, free-to-consumer business models. These include restaurant and hotel comparison sites (generally powered by user reviews). Some open-access and seemingly reliable specialist comparison sites have emerged, often evolving from special interest publications or programming (eg, What Car?).

These sites and platforms have both market and public value. When they do their jobs well, they keep markets clean and can help power the growth of new and challenger products. However, there are many markets within which such intermediaries do not exist, or those that are there have a conflict of interest. Help is at hand to find a good hotel for you and your family to stay in, but what is the equivalent for finding a care home for your elderly parent?

One important question is the market structure around the tech giants, including convergence and conflicts. Amazon has evolved from a specialist provider (of books, etc) to a general-purpose market. Others, such as sports goods suppliers through to hardware stores, are seeking to emulate the Amazon arc, adding onto their websites a mixture of their own products and those of independent suppliers. The potential conflicts of interest are obvious, and regulators will need to be very tight on watching these, with the working hypothesis that there needs to be a separation between platform players and suppliers.

Google, and to some extent social media platforms, are evolving into a very interesting but potentially conflicted space with respect to market operation and design. Paid-for advertising and ranking are increasingly sitting alongside quality and price ratings (often user co-produced). The relationship between these ranking systems is absolutely key to effective market functioning.

This partly reflects human capacity. Few people will scroll past the first, or perhaps the second page of ranking. This leaves the consumer, and platform designer, with a dilemma about how to prioritise rankings. Competing rankings could include:

- Geographical distance
- Consumer or independent rankings of price
- Consumer or independent rankings of quality, which themselves might have multiple dimensions
- Business paid rank order (or size) – itself on the basis of projected consumer value or interest
- Business-derived data on price or quality
- Platform-based estimated matching of consumer preference, product relevance, and value to the platform.

Platforms deriving revenue from advertising have clear incentives to give a higher profile to paid partners. But they have to balance that at least somewhat by consideration of relevance and (increasingly) consumer-derived reports of quality.

An illustration of the dilemmas comes from one of the major restaurant intermediary sites. It faced a decision about what to do with food safety data. Although two-way facing (consumers and business), the platform's revenue came from the restaurants. Though consumers might want to know that a restaurant had poor food hygiene, the paying restaurants did not want this data showing. One compromise, pursued by some platforms, was to have the data on the site, but relatively 'buried' (cf shrouding). In the end, one major platform decided that it would take the hit on advertising revenue by dropping the profile or delisting restaurants with the worst 5 percent of food safety ratings. It did this on the basis that this was a hit worth taking to maintain consumer confidence and its own brand. But was 5 percent the right level? Would it be better to keep the listing, but have the food safety rating more prominent?

Another BIT trial simulated alternative ways of adjusting rankings for electrical goods according to consumer or business interest. It showed large differences in products bought, consumer detriment and business profit. Similarly, data and privacy have become a key battleground with extensive evidence of active shrouding.

A ranking platform that blatantly prioritises business profit (or paid-for ranking) over quality or price will suffer. Some consumers will simply choose not to make a purchase because the product doesn't match their requirements, reducing sales volumes. Secondarily, the platform itself may lose traffic and volume to challenger platforms that offer better guidance, though note that this will be subject to exactly the same constraints as illustrated by the earlier Laibson and Gabaix work. This equilibrium is likely to be skewed by exploitation by firms of naive consumers, some exploitation of firms by sophisticated consumers (eg, always look for the best-ranked firms that don't pay for advertising), and now also firms are being exploited by the platforms.

Tripadvisor/Yelp everything?

There is compelling evidence that markets built around extensive consumer reviews can change and improve market functioning. Harvard researcher Mike Luca found that getting an extra star on consumer ratings causally boosted the revenue of that restaurant by around 7 percent in the year following. Further findings include that the boost is especially large in areas with more tourists; to one-offs (versus chains); and is roughly twice as large a boost to good restaurants as the negative effect on poorly rated restaurants.

The net result is an increase in quality and an increase in market volume: in effect, people become more confident and able to eat out, and a virtuous circle of quality and growth is established.

This evidence illustrates the powerful case for harnessing such feedback mechanisms across as many markets as possible. Indeed, one estimate is that such 'deshrouding' across consumer and B2B markets could roughly double productivity growth.

However, it should be noted that there will be many markets within which consumer feedback alone will not be sufficient, or a reliable method of deshrouding. For example, consumers may rate their financial advisers as charming or helpful, but that may have little bearing on their actual performance. Indeed, previous work has shown that there was no relationship between levels of fund fees and performance, despite consumer willingness to pay (suggesting that consumers wrongly attributed the higher fees and perhaps better treatment as an indicator of performance).

As such, while 'Tripadvisor everything' is a sensible starting point for many consumables, it is likely to need supplementing with other data on a sector-by-sector basis.

Implications

Addressing shrouding and other behaviourally-based market failures offers a way to boost growth, productivity and reduce consumer detriment. In effect, it is a form of enhanced competition policy, to move us closer to something approximating a

broader version of the efficient market hypothesis (see Halpern et al, 2024 for a more detailed discussion).

The creation of the Competition and Markets Authority (CMA) was intended to address these issues by moving beyond the institutional and conceptual models of the Competition Commission and Office of Fair Trading. In short, the CMA was born of a recognition that even seemingly competitive markets will often find a non-optimum equilibrium with considerable consumer detriment.

The CMA and sister regulators should continue drives into addressing bad practices, such as drip pricing, fake reviews, and other sludge. But we need to go further into mapping the scale of detriment and shrouding across sectors, and engage in active market design to address issues at source. Failing to do so creates a risk of a 'whack-a-mole' dynamic, where companies innovate around deception and sludge, rather than competing on true quality and price.

In contrast, the Financial Conduct Authority's (FCA) 'Consumer Duty' is a higher-level, principle-based approach. Under the Duty, firms should be open and honest, avoid causing foreseeable harm, and support customers to pursue their financial goals.[8] Markets and companies are continuing to adjust to the new duty, but it has the major advantage of side-stepping the details of any particular mechanism and instead maintains a focus on the core outcomes for consumers. If it is found to work for financial services, there will be a strong case for extending the approach into other markets.

One strategic approach could be to map detriments, sludge and shrouding sector by sector, and then agree on market changes sector by sector. This process should seek sector and consumer engagement around what constitutes fair assessments of product or service quality, including what data exists or could be collated to enable better consumer or B2B judgements. The test should be: can a reasonable consumer distinguish between products or services by price and quality?

Second, enhancing consumers' – or their trusted agents' – ability to access their own consumption data would provide a complementary step. The original midata act (2011) was quite limited in its scope, though it has broadened since. It enables consumers (and good companies) to make more individually tailored judgements. It opens the door to vendor management by consumers, to match consumer management by vendors. Actions may also be needed to address David vs. Goliath

legal actions by large data holders and companies against small start-ups (cf Rodeo), albeit while balancing the need for adequate data protection.

Third, regulators, governments and other bodies should be prepared to be more active with respect to nurturing and policing market intermediaries, given the crucial and increasingly important role they can play. Where such platforms exist, regulators, governments and media have a key role to play in keeping these platforms clean, such as through mystery shopping and independent assessment of the ratings provided. Being stricter on the line between holding platform status (for all vendors) versus being suppliers in their own right may need to be more strictly enforced.

The emergence of various forms of digital ID platforms offers a likely channel for the emergence of data intermediaries acting on the consumer's behalf. These may prove key to making a reality of the midata powers, and could provide a powerful counterpoint to existing platforms.

Where choices and sectors exist without some intermediaries, regulators and governments should look to create them. That does not necessarily mean that those bodies need to deliver these themselves. A better model will likely be to run competitions with a mixture of funding and data access, starting with sectors with the largest detriment and informational failings, to foster such intermediaries into place. Indeed, independent digital identifiers may be a crucial area for state action, given the role of the state in 'gold-standard' ID verification, in the form of passports, driving licences, birth certificates and so on.

A key challenge for regulators and governments is the structural design and role of the tech giants. Large platforms, and especially those characterised by strong network effects (or close to monopoly status), have deep tensions between their commercial and advertising interests and their capacity to rank by price and quality. Traditional antitrust measures, ie, breaking up such platforms, may have considerable downsides if they result in fragmented markets, destroy real 'network effects', or undermine the ability of consumers to take 'whole of market' views. Requiring such platforms to move to wholesaler status – or 'pure' platforms – may sometimes provide a better way forward, such as how telecoms markets are sometimes regulated to require national infrastructure providers ('wholesalers') to provide access to a variety of competing providers that use this infrastructure. For

tech giants, this could involve considerable restructuring of tariffs, such as between 'wholesale' infrastructure providers and B2C providers, to improve market function and maximise public and private goods.

In such a near-future world, search engines and switching sites would likely derive much less revenue from advertising and paid-for rankings might even be banned (or at least outside very clearly demarcated boundaries. Indeed, in such a world, advertising per se becomes much less important to firms (and consumers) outside the promotion and highlighting of new-to-market innovations or products. Firms instead have to compete for their rankings on quality and price metrics.

Revenues for the (wholesale) platforms would now come from three sources. Tariffs from sales, noting these would now be driven more by robust quality and price ratings. More limited advertising focused on innovation or new product release (note: this might include a shift of revenue from conventional advertising into support for early adopters and independent ratings). Revenue from specialist intermediary platforms that work on behalf of consumers.

The prize is large. Behaviourally-informed markets that are closer to the efficient market hypothesis will be good for consumers and businesses. Good businesses and products will grow faster. Labour and capital will be more effectively directed to higher productivity and value areas, raising earnings and returns. Consumer detriment will be reduced. Resources can shift away from zero-sum advertising to innovation, efficiency, and quality improvement.

There will be an intensification of interest and focus on exactly what drives quality or value (see section on preferences below). For example, there is a live debate as to whether the switching sites that now dominate insurance markets have become too skewed to price only. More fundamentally, there is an appropriate debate to be had as to how quality is measured in many areas. How do we rate the immediate pleasure a food provides compared to how healthy or sustainable it is? Consumers will often be 'satisficers' on one attribute (eg, food safety – 'I'll buy it as long as it won't make me sick'), while 'maximisers' on another (eg, taste – 'I'll pay for better'). Consumers will also differ on their rankings of rankings, such as how much weight they give to the environment, or the trade-offs they make between them.

Finally, many of these 'platform' design issues may be largely overtaken by the emergence of AI assistants. It looks increasingly likely that the digital search activity

of the last decade will be replaced by AI engines that will interpret and bundle our search activity, and will filter and make trade-offs on our behalf. Instead of having to scroll through a ranking of restaurants in a particular place, your AI assistant will be able to filter for you according to the kind of food you (and your fellow diners) like, the available bookings, price and so on. The potential gains for consumers and providers are huge, but so too are the opportunities for conflicts of interest and new and subtle forms of shrouding. It will certainly matter greatly on whose behalf the AI agent is optimising.

But within a better-informed and shaped model, these are puzzles and tradeoffs that markets can solve. We just need to make sure it is a market design based on humans, not imaginary 'econs'.

5. Preference inconsistency

> *Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog. …Give me that which I want, and you shall have this which you want, is the meaning of every such offer…*
>
> Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*
>
> *The great source of both the misery and disorders of human life, seems to arise from over-rating the difference between one permanent situation and another...*
>
> Adam Smith, *The Theory of Moral Sentiments*

A beautiful core principle of economics is that for a transaction to be agreed, it must increase the utility or well-being of both parties. Why else would they agree?

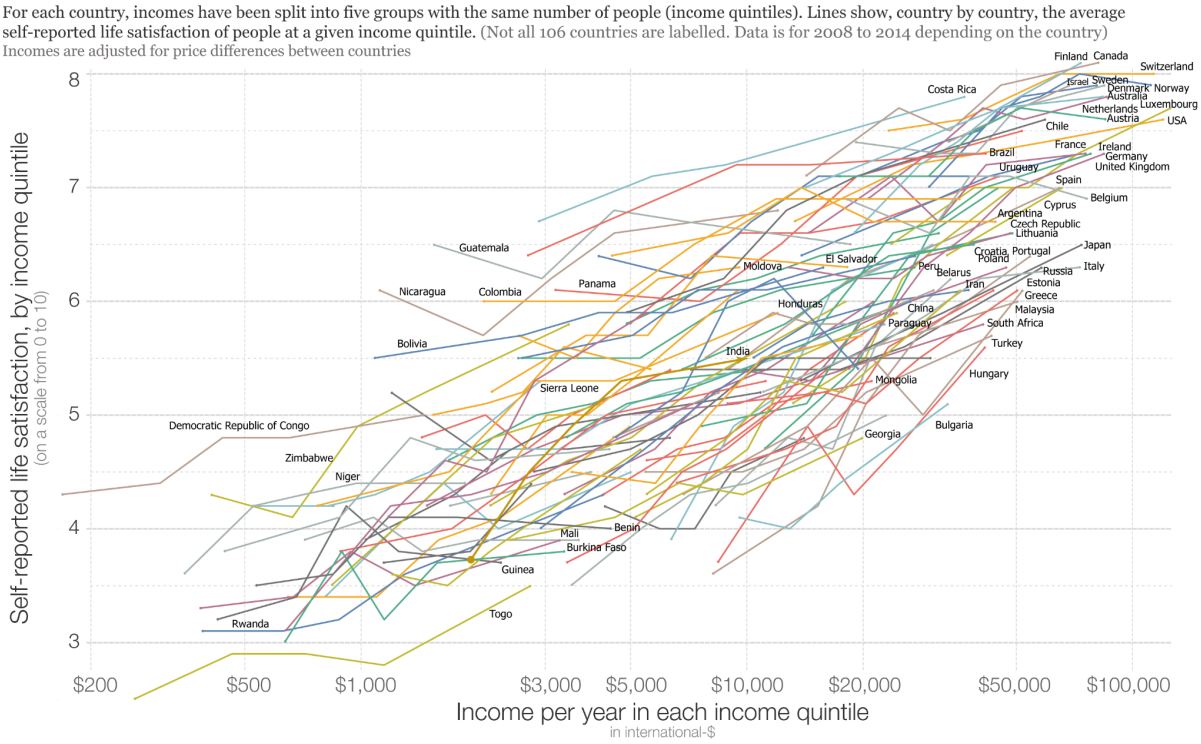
However, a major complication is introduced if individuals' preferences are inconsistent or incoherent. Several classes of inconsistency are possible, where preferences can appear to flip-flop depending on cues or prompts that ought to be irrelevant to the choice. These include temporal inconsistencies, where our preference changes according to the timeframe; and inconsistencies where we simultaneously hold contradictory views on the same action or target, that are tipped according to seemingly irrelevant aspects of the choice architecture. Similarly, we can be inconsistent across our preferences, where a choice appears incompatible or contradictory with another choice already made.

Preference inconsistency creates the possibility that our choices may zig-zag like a drunken walk; that we can be baffled by our choices; and that, at the aggregate level, we may experience a form of '[joyless growth]'.

The stock response is to note that there is no serious alternative to taking a person's expressed preference in their revealed behaviour as the gold-standard guide to their 'actual' preference. Should we defer choices to a super-smart AI or a clever state that will better anticipate what will make us happy? Even if humans have regrets, they can also learn. Finally, there is encouraging evidence that rising wealth genuinely does bring greater well-being, whether measured 'affectively' (how

happy you feel) or evaluatively (how satisfied are you with your life?) This broad relationship between wealth and well-being is found at both the individual and collective levels.

Figure 1: Self-reported life satisfaction across the income distribution, country by country



Credit: Our World in Data

Nonetheless, despite the comfort from finding that money does seem to buy happiness, we need to consider how serious the anomalies around preference inconsistency are.

Evidence

It is now well-established that humans are often inconsistent. When people are asked what movie they would like to watch next week, they tend to choose something highbrow (cf Schindler's List). In contrast, when asked what they want to watch this evening, they choose an action flick (eg, an Arnie or Marvel movie). So what happens when next week arrives? They wish they'd chosen the action flick.

Field and lab studies strongly confirm these effects. For example, in a classic study, people were asked to choose between a healthy or unhealthy snack – fruit or chocolate – that would be delivered in a week's time. Impressively, around three-quarters chose the healthy fruit. But when the day arrived, and subjects were again asked what they wanted, three-quarters now chose the chocolate. Subjects weren't even consistent on the same day. Asked after lunch, people generally chose the fruit, but asked late afternoon, when they were hungrier, most people chose the chocolate.

Most people intend to exercise more and save more tomorrow. The trip to the gym, or the diet, always starts tomorrow, but tomorrow never comes. Financial behaviour shows a similar effect: we genuinely intend to save more for the future, but today's expenditure is often just too tempting to resist.[9] This inconsistency comes in many forms. For example, we value something more after it has become ours.[10]

Scans of human brains help fill out the effect. If prompted to imagine getting injured, large parts of our brain light up as if in pain. But if asked to think of our future self getting hurt, our reaction is greatly dampened. In a powerful sense, our future self isn't us – it's a different person.

An arguably even more serious inconsistency arises within our current preferences, partly depending on whether the decision is made by our fast or slow brain. A wide variety of such effects have been found, including from 'framing effects',[11] context dependence, including the 'decoy effect',[12] and other 'preference reversals'.[13] A current everyday example is most people's revealed preference to scroll for hours on digital devices. Yet when even a small friction is introduced, such as a three-second delay on your mobile, this leads to more than a halving of time spent on such apps. The former is a fast brain ('system 1') preference, while the latter is a slow brain ('system 2') preference. Which is the real you?

A further example is manifest in 'goods that people but wish did not exist'. Social media has attracted much attention in this respect. When asked individually, people say that they would have to be paid a considerable sum to give up social media. But they also say they would pay a sizable sum to live in a world in which no one used social media.

Clearly, such goods and preferences embed an equilibrium issue, and in that sense, the seeming inconsistency is perfectly rational. If all your friends are on social media,

you would be missing out if you stopped using it. But if you all left the platform together to talk to each other or go down to the pub, that might be even better. It's just that you have a big coordination problem to achieve it (especially if that network is most of the world).

This preference asymmetry is not restricted to social media. In his classic text, Tibor Scitovsky highlighted a number of examples of the 'joyless economy', or what we would now call joyless growth, arising from preference inconsistency. For example, he noted how individuals might choose cheap mass-produced furniture, but then bemoan that craft, hand-made furniture was no longer available (the craftsmen and women having been put out of business).

The psychologist Danny Gilbert has documented multiple cognitive biases that lead to these misprediction errors. We systematically misremember what made us happy, and in turn mispredict what will make us happy.

Even more alarming, evidence suggests that when subjects' own choices are (secretly) reversed, they still go on to explain why they prefer the choice that they didn't make. This work benefits from the unusual talents of magician-turned-psychologist Petter Johansson. His experiments hinged on asking subjects to choose which face was more attractive, or answering questions on their preference between policies and political parties, and then asking them to explain their choice. The twist is that the subject was actually given the picture, or policy view, that was the opposite of the one they chose. Subjects would rarely notice the shift and proceed to offer an account of their preference just the same. This artificially altered feedback on the person's own views also led to lasting changes in the person's 'actual' views.

So-called 'cognitive dissonance effects' have a similar dynamic to 'choice blindness' effects, without the sleight of hand. In a famous series of experiments, Festinger et al paid subjects to take part in a boring task and to persuade someone else to take part afterwards. Paradoxically, subjects who were (randomly) paid less reported the task as more interesting and were more persuasive in getting others to do the task. Festinger argued that the poorly paid subjects resolved the 'dissonance' between the poor pay and their own behaviour (having done the task) by changing their perception of the task: 'I obviously can't have done it for the money, so it must have been fun.'

A final class of preference inconsistency arises when we move from the individual to the group level. Hearing or seeing the preferences of others often leads to a shift in our own behaviour or preferences. This can be relatively automatic (or 'system 1' mediated), such as our strong tendency to want something if we think others want it, or not to litter if we think others aren't – so-called 'social proof'. It can also be more reflective, such as in a deliberative process when people seek to reach a consensus about what to do on climate change, obesity, or even how much damages to award. The results of such processes are often dramatically different from the sum of the individuals' expressed preferences prior. For example, deliberative polling of the UK public found people shifted from a majority in favour of more punitive punishments to a majority against, while in the US, deliberative polls have been found to move ahead of subsequent shifts in public attitudes to same-sex marriage. Deliberative processes also help people converge on consensus positions and reduce the gulf between extreme polarised views often expressed in conventional polling, such as shown in a wide-ranging deliberation with 500 Americans in 2019 – America in One Room.

At least part of the dynamics of deliberative processes reflects people updating their priors about the impacts on others, and as such being good Bayesian (and good humans – cf section 5 below). However, it should be noted that other circumstances can prompt 'groupthink' dynamics, where individuals in a group adopt much more extreme positions than the average of their personal views, notably as a result of social influence and a tendency to seek confirming evidence.

Implications

The implications of preference inconsistencies and reversals are profound. They are a puzzle for any human to know themselves.

For the policymaker, they destabilise core elements of the 'Green Book' method, through which cost-benefit analyses should be employed to help decide which policy should be favoured. Willingness to pay is one such technique, but we can now see that it may be highly malleable, or even show sign reversals (flipping from a willingness to pay towards a requirement to be paid, according to how the question is asked).

One option may be to capture a wide range of preference methods, and to express the answer as a distribution or range. Distributions will still differ, even if they sometimes overlap. For example, loss of life will still generally be valued more than a minor inconvenience.

Another option is to take a view on whether some forms of preference should be given more weight than others. Most obviously, should a reflective (system 2) preference be given more weight than a rapid or impulsive (system 1) preference? We clearly make this distinction in some domains of life. A law court is likely to put more weight on a contract entered into following careful consideration by a purchaser or seller than one that was entered into without reflection or awareness of its details. Both judges and juries are inclined to give higher sentences for premeditated murder than for manslaughter. Relatedly, people frequently voluntarily seek to constrain their impulsive selves by entering into some form of commitment device set while in a reflective mode, such as putting savings into an account that restricts access.

Insofar as our preferences are interdependent, this becomes a matter of governance design or democracy. Given that preferences are inconsistent, it is likely that the design of a collective decision-making mechanism will itself affect the outcome. A 'minor' example is how voting preference may be affected by the environment within which the ballot is cast, such as being more likely to vote for conservative policies when symbols of death or religion are present (such as in a church) and more likely to vote for left wing policies or support of public goods in places such as schools or fire stations.[14]

A more serious example is how the extent of reflection built into a deliberative mechanism may affect the outcome. For example, deliberative mechanisms that require participants to spend many hours studying a particular issue and talking with other participants will lead to systematically different preferences coming to the fore than a voting mechanism that requires very little prior thought or that bundles large numbers of issues into a single decision. The former will be lodged firmly in system 2 decision making, while the latter will likely be much more affected by system 1 heuristics. A recent example is how UK citizens, given the time and opportunity to deliberate, converged on a much softer version of 'Brexit' than that expressed in polls, or indeed the formal political process. An even more current example is the finding that a deliberative forum on the use of AI in controversial areas of public

administration, such as prisoner release or worker leave through ill-health, showed significant shifts towards more support for AI (itself largely mediated by increased knowledge).

One design principle that might be used is the reliability and robustness of the decision-making mechanism. A decision-making mechanism could be judged unreliable if modest and seemingly irrelevant variations in starting conditions or process lead to a different result. In contrast, a mechanism that is robust to such variations might be favoured. This ought to be a testable proposition. Indeed, a recent large-scale deliberative process conducted for meta, conducted by the Stanford Lab and the Behavioural Insights Team, across multiple countries, in effect did this. Focusing on what the rules should be about AI behaviour in the metaverse, the results showed a blend of conclusions that were similar within and across countries, while other conclusions differed across cultural contexts (for example, deliberative groups within Spain and Brazil felt that AI-human romantic relationships should be restricted). We are familiar with asking about the reliability and robustness of many mechanisms, from medical tests to recruitment processes. We should ask similar questions about the decision-making processes within our political economy.

This is not an abstract matter. Finance Ministers and Governments feel genuinely constrained – and probably should be – by the democratic processes and commitments they made within them. Political parties regularly make commitments within electoral campaigns that their officials at least regret, such as not raising certain taxes. Yet one can readily conceive of a situation where such tax decisions could be guided by much more in-depth public deliberations around budgets, rather than bundled into the single vote at a general election.

For example, Texas ran a robustly constructed deliberative forum on energy policy and taxes in the late 1990s that (unexpectedly) came out very strongly in favour of doing more on renewables, which in turn changed policy and led directly to a significant and sustained investment in renewables. An arguably stronger form is 'participatory budgeting', within which citizens are given real responsibility to co-decide budgetary and tax decisions. Originally pioneered in Latin America, and most famously in Porto Alegre, such processes can increase citizens' support for the provision and funding of collective goods, and are a powerful form of genuine citizen voice – a more considered and robust expression of preference.

Finally, we might note that preference consistency might be helped by 'boosts' that enhance the decision-making capabilities of people and markets, at least under conditions of heavy shrouding (see above). This has an extra edge given the likely emergence of AI decision aids. Which type of preference should I set up my decision aid to guide me towards? Do I instruct it to maximise my short-term hedonic gains (immediate pleasure), my longer-term life satisfaction, or collective well-being?

A likely element of a sophisticated decision aid is that it will need to ask you, presumably in a reflective moment, how you wish to balance between these rival preferences. In this sense, it will echo the questions that a good financial adviser might currently ask about your appetite for risk, the trade-off between short-term and long-term gains, and your wider goals.

6. Social capital – irrationally cooperative?

> *Unfortunately, many analysts – in academia, special-interest groups, governments, and the press – still presume that common-pool problems are all dilemmas in which the participants themselves cannot avoid producing suboptimal results, and in some cases disastrous results.*
>
> Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action*

Economists have long been bothered by a 'problem' that, for most people, seems like a good thing. Humans are a lot more cooperative than economic theory implies. You might call it 'nice'.

This has been illustrated by a series of widely repeated experiments. These involve situations where people have a choice about how to engage with others. Perhaps the most famous is the 'ultimatum game' within which one player gets to decide how a pot of money is divided, and the other gets to accept or reject this 'offer' for them both. Economic theory suggests that the player dividing the pot should award themselves the vast majority of the money (say, 99c in a $1 pot). The other player should then accept the offer, since to reject would result in them losing money (even if it was just 1c).

Yet that is not what happens. Instead, most people make an offer much closer to 50:50, and most 'deciders' reject offers that give them less than around ¼. In short, humans are a lot more cooperative – or generous-spirited – than classical economics predicts they will be.

Several of the 'anomalies' pulled together by Thaler in his essays, and later in The Winner's Curse, concern this troubling violation of 'rationality'. These include a chapter on the puzzle of 'Cooperation' and a whole chapter on 'The Ultimatum Game'. Excessive cooperation appears to be widespread. I recall being briefed by a frustrated economics PhD student on how his experiments 'weren't working'. He had created a series of 'games' to see how much people would cheat on their (within the experiment) 'taxes'. The problem was, he couldn't get his participants to cheat. Instead, they dutifully kept paying.

To people outside the economics profession, this puzzlement is puzzling. Across most of the world, people walk into experiments with strong notions and 'habits' of fairness. They mostly bring these notions or habits into their interactions with strangers, not just into interactions with those they already know or are in a 'repeat game' with.

Evidence

The ultimatum game, as mentioned above, has been widely studied. A meta-analysis across 75 studies found that, on averag,e proposers offered around 40%, ie, slightly less than half, but way above the 'rational' 1% suggested by economic theory. On average, such offers were accepted in more than 80% of cases. Responders were more likely to reject lower offers, though more forgiving if the overall pot was larger. Very limited differences have been found in proposer offers across countries or cultures, though some differences have been found in what the responders would accept.

Curiously, one of the most striking outliers to this pattern was found by Henrich. He replicated the Ultimatum Game in the Machiguenga, a relatively isolated tribe that lives in the tropical forests of the southeastern Peruvian Amazon. Henrich found the average offer was only 26%, and the modal offer just 15% (contrast this to a modal offer of close to 50% in many Western countries). Despite these low offers, Henrich found that the vast majority were accepted.

It is rather ironic that the one place in the world where people were found to behave along the lines of classical economic models was where nobody knew about economics or Western market economies. Henrich concluded that notions of fairness that pervade behaviour in most Western countries derive from cultural learning.

In fact, this cooperation has enormous economic benefits. An everyday example is provided by Ronald Coase on 'The Nature of the Firm' (1937) – or why do firms exist? Coase concluded that the central benefit of firms, and other such organisations, is how they lower transaction costs (as we would now call them). In essence, firms create a bubble of trust within which people cooperate and ask things of each other, without continually having to contract and negotiate. Coase was later awarded the Nobel Prize in Economic Sciences.

In a parallel literature, social scientists have called this asset, or cultural characteristic, 'social capital'. Since Robert Putnam onwards, most scholars have defined this social capital as social networks and trust. In crude terms, social capital is the people that you know, and your ability to trust them to cooperate or to help you when in need.

In contrast to the lab results for the ultimatum game, social capital is found to vary substantially within and between countries. A central measure – ie, that aligns closely to a principal component analysis of a range of social capital measures – is a simple and widely used survey question: 'generally speaking, do you think that other people can be trusted?' People are asked to respond either 'yes' or 'no, you can't be too careful' (or 'don't know').

The proportion of people saying 'most people can be trusted' ranges from around 60-70 percent in Nordic nations, to less than 10 percent in much of Latin America and Africa. This 'social trust' also varies markedly within countries. For example, in Anglo-Saxon countries, people who are older and those with higher levels of education are much more likely to report that others can be trusted. In this sense, professionals in their 50s living in the UK or the US report trust levels akin to those in Sweden or Norway. On the other hand, younger people with lower education levels report trust more akin to southern Europe or Latin America. It is as if they are living in different countries.

These differences appear to be highly consequential. High-trust countries economically grow at faster rates. This appears to be mediated by a number of factors, including better information flows (cf shrouding above); lower transaction costs (being able to do a deal on a handshake), and better labour market functioning (employing the best person for the job instead of your cousin).

Higher social capital is also predictive of higher earnings at the individual level. Having lots of well-connected people on your WhatsApp, whom you trust, is great for your career prospects (cf Granovetter, weak ties). Indeed, recent US work suggests that a young person's 'bridging social capital' or 'economic connectedness' is a better predictor than background, race, or where you live.

Interestingly, it helps to be trusting in a well-calibrated way. Individuals who express a level of social trust that is much higher, or much lower, than that which is prevalent in

their society earn less than those who are better calibrated. This drives home the point that this is not about 'blind trust', but trust based on actual trustworthiness.

Implications

The most basic of policy implications is to measure and systematically value social capital. We are now relatively familiar with how to measure other forms of capital: financial, physical, environmental, and even human. Treasuries, in turn, plug these into growth and other models. Why not social capital, not least since it appears to have more predictive power than some of these other capitals?

The failure to measure or factor in social capital distorts both our accounting and policy. A joint project between Prime Minister Blair's Strategy Unit and the Office for National Statistics concluded that the rough 'direct' value of social capital, notably expressed through informal care, reciprocity and volunteering, was around 50 percent on top of existing GDP.[15] But failing to measure such a huge sum, or factor in changes within it associated with policy changes, can make a nonsense of any serious cost-benefit analysis. Similarly, shifting millions of children into paid-for childcare from informal care can appear as a massive boost to GDP as a large 'off-balance sheet' activity is moved into a costed activity, regardless of whether it was better or worse.

One argument deployed against worrying about social capital is that, even if you buy the argument that it is important, there is little that governments can do about it. For example, Putnam roots the dramatic differences in social capital between the north (high) and south (low) of Italy in the repeated invasions that the south experienced over a thousand years of history. What then can a policymaker hope to do?

Similarly, variations in levels of social capital in the US today broadly map onto social capital levels found in the countries from where people's ancestors migrated generations before. Why are Minnesota and the Twin Cities so high in social capital? Lots of high-trust Swedes moved there a hundred years ago, and appear to have passed on those habits of trust across the generations (including mediated through rich civic and state institutions).

Yet there is plenty of evidence that shorter-run changes can make a difference too. Rothstein, for example, suggests that the relatively high social trust seen in Sweden

today is rooted in changes that occurred in the nineteenth century, and particularly the spread of legal practices based on impartiality of treatment across social strata. It implies that it was a relatively rapid shift that drove this higher social trust.

Even more extreme, when subjects engage in a lab experiment where higher returns are achieved through cooperation (such as the prisoner's dilemma), they end up much more likely to report that 'most other people can be trusted'.

Haldane and Halpern (2025) suggest a range of policy interventions to boost social capital and cooperation. At the micro-level, this might focus on building social skills and habits of meaningful connection. Examples include school-based programs that nurture relationship skills and cooperation; out-of-school activities and clubs, particularly those that create 'bridging social capital' between kids and adults from different backgrounds; and residential higher and further educational experiences. On the latter, it is noteworthy that young people who go to university generally end up with two to three times the level of social trust as those who don't go to university. It would be good to have stronger studies to pull apart the causality, but it seems that a substantial part of the earnings and other premiums that flow from higher education stem from the social capital built from bringing young people together in a safe environment to build new relationships, and the habits of trust and trustworthiness that then last a lifetime.

At the meso level, interventions include the creation of spaces and places where people can meet and connect. Most obviously, these are community centres, parks and common spaces that enable people to connect and socialise. These spaces create opportunities for people to connect, but do not force them to. It seems that the extent to which such spaces create the opportunity for easy connection, but not forced interaction (cf a crowded train), is key. In the words of Frost, 'good fences make good neighbours'. Well-designed commercial spaces can also play a role. For example, the Starbucks 'Urban Coffee Opportunities' programme – prompted by Magic Johnson – led to the setting up of Starbucks cafes in 100 poor neighbourhoods where the coffee giant would not normally have opened. Follow-up studies showed that these neighbourhoods went on to have double the business start-up rates than matched control areas over the 7 years following (roughly 6 versus 3 firms a year). Having a space where people could meet and work appeared to boost the opportunity for connection and cooperation, which in turn boosted small company formation.

One potentially powerful but controversial intervention at the meso level is to deliberately induce more social mixing at the school or even classroom level. The strength of the effects found by Chetty in the US – and largely replicated in the UK – has reignited this debate. A more oblique approach might be to seek smaller educational institutions, on the basis that (ironically) more social mixing tends to occur in smaller schools and colleges. This seems to occur because smaller institutions provide less room for internal sorting into social, religious and ethnic cliques. UK data also suggests a potentially big role for 'hobby' and sports activities, such as five-a-side football, made easier in the UK than US due to a smaller geographical area – the rich and poor still live close together in the UK and much of Europe.

At the macro level, interventions might include investing in cultural and cultural events that create shared cultural experiences; avoiding welfare and other systems that inadvertently signal distrust; and legal and market mechanisms that punish untrustworthy behaviour. This last point overlaps with the importance of 'deshrouding'. Digital networks and feedback mechanisms can have a key role in making it more costly for bad actors to cheat, in effect shifting the equilibrium towards higher trust.

One can argue that some of the latter types of interventions – such as well-policed feedback mechanisms and digital reputations – are on a fine line between nurturing social capital and substituting for it. Arguably, they do both. They make it possible for relative strangers to cooperate and trust each other.

But that is a key element of human advantage. Young humans do no better than chimps on many cognitive tasks. The one area where humans massively outperform our primate cousins is how rapidly we learn from each other. Chimps are highly social within their troops of less than 200, but will typically violently attack or avoid those from other troops. Humans generally do not. Social capital and cooperation are a massive asset. We should celebrate that classical economic models are wrong on this point, and look to leverage it in more sophisticated models and policy.

7. Institutions

> *Government has three primary functions. It should provide for military defense of the nation. It should enforce contracts between individuals. It should protect citizens from crimes against themselves or their property. When government – in pursuit of good intentions – tries to rearrange the economy, legislate morality, or help special interests, the costs come in inefficiency, lack of motivation, and loss of freedom. Government should be a referee, not an active player.*
>
> Milton Friedman, *Capitalism and Freedom*

Even in neo-classical Chicago School models, there is room for institutions. These include legal systems to enforce contracts, antitrust measures to prevent monopolies, and governments to organise the provision of basic public goods, such as roads, basic education, and national security.

Nonetheless, such models presume that most coordination is best left to the market. Prices will move according to supply and demand, enabling an emergent web of coordination and allocation that will exceed anything that a state or formally coordinating body can achieve. This is an elegant organising heuristic for policymakers in many Western market economies, not just as a playbook for economic growth, but as protection of core freedoms for citizens.

Indeed, there remain major areas of policy where the argument can be made that an expansion of market freedom could be a really good idea. Top of the list in the UK might be housing and planning, where successive planning constraints have essentially crushed the relationship between house values and development, such that around 4 million too few houses have been built and those that have been built are not where people (or market prices) suggest they are most wanted.

Planning aside, the shift away from state management of the economy was central to the Thatcher and Reagan reforms. This included reducing state subsidies, reducing taxation and in the UK, a major program of privatisations of state industries.

The embodiment of this belief in the thinking of officials, and as expressed in policy detail, runs particularly deep in the UK – perhaps even deeper than in the US. As one senior official privately remarked (during the Cameron-Osborn era): 'The Treasury

knows we have to have a business department for political reasons, but given a free hand, they would shut it tomorrow. They think the best thing that Government can do for business is just get out of the way.'

Within the UK policy community, industrial policy and market intervention became instinctively frowned on. The unease about 'industrial policy' persisted through the Blair and Brown Administrations and through into the Coalition and Conservative Administrations of the 2020s. You just needed to utter the words 'British Leyland' – the failed car manufacturer that received around $16 billion of state subsidies – to close off the argument. Oliver Letwin, the Conservative Minister who was Head of Policy in the Cameron Government, made the comparison to Stalin's disastrous 5 year plans. His point was that markets were much better at solving the complex allocation and incentive problems involved in making tractors than state institutions, let alone allocating resources within the complexity of a contemporary Western economy.

Frequent reference is made in Whitehall to the way in which economically successful countries over recent decades have pursued stripped-down, free-market economic policies. It has been argued that the UK should become 'Singapore-on-Thames'. Yet there is a strong case that UK governments have gone too far in the belief that the best thing they can do is 'get out of the way'.

Leaving aside the point made by Singaporean policy makers themselves – that if you just took Greater London, its economic performance looks very much like Singapore – a closer analysis of economic policy in countries like Singapore and the US doesn't fit the caricature of it in Whitehall.

Evidence

Higher performing nations, in terms of their economic performance, generally have intermediate institutions to play a coordination function. Perhaps the most obvious contrast is between the outstanding successes of the Asian 'tiger' economies relative to other regions and countries.

A key characteristic of the economic strategy of the Asian tigers was an overt industrial strategy to move up the value chain. This was twinned with sustained and deep engagement with industries identified as of 'strategic importance' to work out what it was that the government needed to do, and what it was that the relevant sector and industry players needed to do. The exact details of the institution(s) that

fostered strategy and compact varied. In some cases, this was done within the government itself (cf Singapore), in others it was done via arms–length bodies such as sector councils or some form of Productivity or Industrial Strategy Council. All have in common the creation of a 'national system of innovation' (Sainsbury, 2013).

In contrast, countries that dutifully followed the classical account of market reforms, such as many Latin American countries, often found themselves with disappointing economic results. Stripping away the barriers that were presumed to be stopping the market from working, such as protectionism, restricted capital markets, and so on, seemed at best to be a necessary but not sufficient condition. Indeed, many researchers now think that even this may not be true, for example, nascent industries may need a period of protection before they are ready to compete on a world market.

An analysis of US economic policy suggests that it pursued a much more activist economic policy than implied by the Chicago model. In particular, the US government has long funded not just primary research, but the active translation of that research into practice. This included supporting multi-billion-dollar government-funded labs for strategically important applied technology (eg, MIT's Lincoln Lab); co-funding industry efforts to develop technology (eg, Apple, Boeing or Tesla); and providing more general support for R&D in firms (eg, R&D tax credits). These active supply-side polices sat alongside more familiar activities such as patent protection, predictable legal systems, and stable financing.

US activism within its own market has continued, including a sharp debate about antitrust measures among the digital giants. As Lina Khan, until recently the head of the FTC, put it in 2024: 'In the early 2000s, there was a sense that digital markets are so fast-moving that if you are to see problems of market power, that will self-correct, and so the best thing for the government to do is to get out of the way.'[16] To Khan, who herself wrote a famous paper on why governments needed to intervene to fight new forms of monopoly, the answer was clear: 'To stay ahead globally, we don't need to protect our monopolies from innovation: we need to protect innovation from monopolies. We need to choose competition over national champions.'

In this sense, US policy combines active market support to build up sectoral capabilities, with active interventions to break up monopolies to maintain market

dynamism. Interestingly, this is actually quite similar to the strategy pursued by China, the US's increasingly strong rival. China identifies 'strategic emerging industries', to which it then provides a range of support. The evidence indicates this strategy has been pivotal to the Chinese economic miracle. This includes promoting tax benefits, firm loans, industrial concentration, and R&D activity – all to increase the supply of technology, talent, and capital. This surge is followed by the stimulation of aggressive competition between firms, driving down costs while maintaining quality to build a comparative advantage both domestically and internationally. The Chinese model is not about supporting single 'national champion' companies (unlike, say, the industrial policies of some EU nations).

The UK, in contrast, ranks highly for primary innovation, but is distinctly underpowered when it comes to transmission and adoption. The UK ranks 5th among 133 nations on the Global Innovation Index 2024 (WIPO), and an impressive 3rd for creative outputs. However, it ranks 14th for business sophistication, 26th for institutions, and 31st for knowledge absorption (and an abysmal 107 for gross capital formation).

It is a longstanding issue. The UK's weakness was laid bare by Andy Haldane, when the Bank of England's Chief Economist, in a speech: 'The UK's Productivity Problem: Hub, No Spokes.' Alongside a broadside of statistics, largely unchanged today, he contrasted the thinness of the UK's economic diffusion and adoption institutions with those of other countries. Germany's 72 Fraunhofer Institutes and the Steinbeis system (providing technical advice to companies) were more than 10 times the size and budget of the UK equivalents, including its 10 Catapults. The German system also twinned the system of advice with three layers of business finance, including the Sparkassen (financing local businesses), the Landesbanken (operating regionally and serving larger companies), and KfW, Germany's national development bank (serving large companies) – altogether far more extensive than the UK's equivalent.

This institutional gap, and the corollary slow diffusion of new technologies and practices, explains much of the productivity gap between the UK, France and Germany. Given its careful documentation, by Haldane and others, it becomes hard to explain the slowness to address the gap among UK policymakers other than as the result of paradigmatic belief in the ability of the market to diffuse and drive adoption by itself – a belief not shared by the policymakers in the UK's more productive neighbours.

The importance of institutions in promoting strong – and inclusive – economic growth has been raised further over the last decade by political economists seeking to explain the differential growth rates between countries. Indeed, this was recognised in the award of the 2024 Nobel Prize in Economics to Acemoglu, Johnson and Robinson. As James Robinson highlighted in his recent visit to the UK, including in a session at No. 10, previous models had largely missed the role played by inclusive institutions as a driver and catalyst of economic growth.

In sum, even the most ardent originators of the Chicago school saw some role for institutions to boost growth. This was limited to a clutch of foundational roles, such as having an independent central bank to maintain a stable currency, patent and legal protections to enforce contracts and provide incentives to invest in intellectual property, along with periodic efforts to break up emergent monopolies. Subsequent cross-national studies, and the practical experiences of trying to boost growth, have shown that governments have an important role in curating and catalysing economies, and in particular by fostering or creating institutions that coordinate state and private sector activity around strategically important sectors of the economy.

Implications

A good summary of the lessons and implications of the role that government institutions need to play to boost economic growth is provided by Rodrik.

> *The hallmark of development is structural change – the process of pulling the economy's resources from traditional low productivity activities to modern high productivity activities. This is far from an automatic process, and requires more than well-functioning markets. It is the responsibility of industrial policy to stimulate investments and entrepreneurship in new activities, especially those in which the economy may end up having comparative advantage. The usual argument against industrial policy is that governments can never pick winners. I show that this is the wrong way of thinking about what industrial policy does. Appropriately structured, industrial policy is a process of strategic collaboration between the private and public sectors, where the objectives are to identify blockages and obstacles to new investments and to design appropriate policies in response.*

> Dani Rodrik, *One economics, many recipes: globalisation, institutions, and economic growth*, 2007. Pp 7-8.

Rodrik is a clear advocate of many of the insights of classical economics, but they need to be supplemented, not least by contextual knowledge.

As the new government came to office in 2024, the UK was in a small minority of OECD nations to lack some form of Productivity Commission or Council to develop, advise and guide its industrial policy. The UK government's capacity to play an active role in the economy is fragmented across a range of institutions and Departments, including its non-statutory National Infrastructure Commission, along with a range of Whitehall Departments (and regional governments).

The first implication is that the UK needs to create a robust coordinating institution to develop and guide economic policy, or strengthen existing ones to play this role. The new Government has now launched an Industrial Strategy Advisory Council, composed of sixteen experts from industry, academia and government. However, this looks remarkably similar to previous short-lived Advisory Councils that have lacked the resources or longevity to make a real difference. The new Council may have an advantage over its recent predecessors in that it has been set up early in the Parliament of a Government with a large majority and a reasonable prospect of staying in power for at least another four years. For this Council to succeed, it will either need to track towards a stronger statutory basis with real resources and leverage over other bodies, or Ministers will need to empower it politically by showing that they will be strongly influenced by it.

Second, studies of other countries suggest that successful industrial policies rest on sectoral specifics, with the construction of detailed mutual commitments between government and industry. There is also a need for a stronger conveyor belt of institutions to accelerate the adoption of sector-relevant technologies and improved management practices. This role is unlikely to be played by a high-level general-purpose single advisory council, however talented its part-time members. As such, the UK is likely to need to develop in-depth strategies for each of its key strategic sectors. These could, in principle, be led by bespoke industrial strategy councils sponsored within government, such as the Office of Life Sciences. However, downsides are that they are vulnerable to high staff and organisational turnover, a lack of deep sectoral expertise, and perhaps most crucially, their embedding in

government means that industry tends to approach with lists of requests and lobbying rather than mutuality and partnership.

A better alternative may be to create or utilise trusted intermediaries tasked with developing in-depth knowledge of the sector, and bridging between industry (incumbent and challengers) and government (local and national). For example, it is plausible that some of the UK Catapults could evolve into this role, though that would require some reworking of their objectives as set by Treasury and Innovate.

Thirdly, governments need to free themselves from the more dogmatic aspects of their attachment to the Chicago model, to enable a genuinely robust and open-minded consideration of the role and design of the institutions that can boost growth, including those that may be absent (or that are present but obstructive).

This does not imply that the state should plan or manage everything: government failures can be as serious as market failures. Instead, it implies a careful consideration of where market dynamics and entrepreneurialism can be harnessed, or even amplified (see, for example, section 4 on shrouding), alongside protection and investment in public goods. There should be room for extensive private development of housing where the demand lies alongside the provision of green spaces and effective transport systems. Similarly, good institutional design is integral to fostering industrial policy that enlarges the space for entrepreneurialism alongside a scaffolding of skills, knowledge and assets to enable industry to thrive.

8. Conclusion

> *Innovation often comes from unexpected sources; we should be open to diverse perspectives and ideas.*
>
> Esther Duflo

This essay has considered some of the key challenges, or 'anomalies', to the classical economic model that remains dominant in the corridors of governments across the world. As Richard Thaler noted at the seminar in Downing Street in late 2024, adjusting classical economics in the light of these challenges may make economic models more complicated, but also move them closer to an approximation of true.

Drawing out the implications should be of critical interest to policymakers. The prize is to uncover new policy levers and upgrade existing levers – or what Dufflo has called getting the 'plumbing' right.

Mental accounting could open the door to more effective taxes and subsidies for the same pound or dollar. Deshrouding the economy offers the prospect of substantial increases in productivity growth, alongside potentially lighter touch approaches to market regulation. Factoring in a more nuanced account of preference, and of cooperation and social capital, points to new levers to boost growth and to lift 'utility' or wellbeing more directly.

A common theme across these areas is a challenge – or refinement – to the claim that markets are the best way of sharing and sifting information, and in turn allocating resources. Sentiment-based decision-making in markets points to emotional reactions displacing information about 'fundamentals'. Shrouding points to information failures – both deliberate and unintended – undermining the ability to make good decisions, slowing the growth of good firms and products. Social capital points to the use of non-market channels for the flow of information and mutual support. The need for intermediate institutions to accelerate the adoption of new technologies and practices points to the ineffectiveness of markets to do this themselves (outside a top tier of 'superstars').

In a comment on an earlier draft of this paper, Cambridge economist Diane Coyle noted that 'a lot of the behavioural insights (and other shifts away from neoclassical

Econ) are by now pretty mainstream in the academic discipline, so why is the policy world stuck with the previous paradigm?' There are certainly brilliant and expert economists within governments, as well as generalist civil servants, who rely on basic models that they have not updated. A common target is to blame politicians, some of whom, in turn, point to a contemporary political discourse that drives over-simplified arguments: hedgehog (simple) views win over fox (complex) accounts in a sound-bite world, even if they are less accurate. My sense is that the reality is a bit messier and human. Even the smartest public servants are prone to an 'illusion of explanatory depth' – familiarity creates a misleading halo of understanding that can be hard to see, let alone update.

This is not a call to abandon economics, let alone markets. Rather, it is a call to develop a more sophisticated account of markets that can lead governments, regulators and innovators to enhance how markets and societies work. A good place to start might be to revisit and sharpen our understanding of public, club and private goods, particularly around information. A first-principles analysis, incorporating behavioural insights, would likely lead us to a very different landscape than the one that we currently have. We live in a world flooded with information, even if much of it is 'junk information'. The marginal cost of reproduction is close to zero. Meanwhile, the distraction economy places increasing strain on the bottleneck of human attention. Much of the value of information and attention, including its public good elements, is increasingly captured by private corporations, and sometimes held but underutilised by public bodies.

These insights and proposed policy levers may not all prove correct or effective. But they are plausible and strong enough to merit serious exploration and robust testing, such as through the new UK Government's 'Test and Learn' program, and through more agile and imaginative approaches to regulatory innovation.

One of the ironies of life in the Treasury is that while it frequently demands better evidence from Departments of State seeking funding (rightly), it has historically been far less robust in the evidence standards it has applied to its own policymaking. Sometimes this has been justified through the claimed market sensitivity of its announcements, notably around tax, such that its decisions can't go out for extensive consultation without altering market behaviour. Yet there are plenty of areas where Treasuries could test policies earlier and better.

In recent years, the UK Treasury has much to be proud of in how it has started to promote more evidence-based approaches to policy and spending, such as through support of the Evaluation Task Force and the fostering of 'What Works Centres' across domestic policy. It should apply the same curiosity, openness and rigour to its own thinking and policy work. The types of policy ideas in this essay should not be dropped wholesale into the next budget or Spending Review. However, the best of the ideas should be tested at local or sectoral levels, or at least in simulated environments.

Dufflo's (Nobel 2019) words are a good challenge to today's economic policymakers. She also notes 'through rigorous evaluation and analysis, we can develop strategies that are truly effective.' To paraphrase, we need policymakers who are daring yet humble. We need daring, in the sense of actively exploring new possibilities and ideas. But we also need to alloy this daring with humility – testing rather than presuming that our ideas will work.

At the time of writing, the world is going through a major economic and geopolitical shift. The Trump administration has introduced tariffs at a level not seen for a century. Geopolitical shifts in power are spilling out with a rising China, an (economically) falling Russia, a stagnating Europe, and a bellicose US. Other nations are having to respond to these shifting forces.

Within the UK, the new Government has started to use its large majority to take on some long-standing weaknesses in its economy, notably to loosen the constraints of its planning system, which has slowed the UK's ability to build infrastructure and housing (the UK ranks 51st on general infrastructure in the WIPO 2024). Yet beyond this, the UK's deep attachment to the neo-classical model runs deep. As Mohammed El-Erian has noted, UK policymakers rely too heavily on the Bank of England and monetary policy in lieu of more fundamental reforms.

If any of the policy ideas thrown up in this paper from updating the neo-classical economic model are shown to work, there will be much to celebrate. If the ideas don't work, then that will be a good challenge back to the 'misbehaving' economists who dare to challenge the classical model. Either way, at least a few Treasury economists will get to celebrate a little more freedom from being 'slaves of some defunct economist'.

Endnotes

1. Includes detailed cost-benefit analysis as set out in the Green Book and its equivalents, and the growing support among Treasuries for robust evaluations, including RCTs.
2. I am indebted to my old colleague Stephen Aldridge for highlighting this point, and for his careful accounts of 'government failures', including in public remarks at LSE and IfG.
3. At the time of writing, Richard Thaler and Alex Imas are working on an updated version: *The Winner's Curse: Behavioral Economics Anomalies Then and Now* (likely publication 2026).
4. In their forthcoming book, Thaler and Imas note that 'despite the empirical robustness and conceptual importance of these findings, insights from behavioural economics remain largely absent from mainstream economics textbooks'. Others are more upbeat, at least about the shifts in the academic community. Diane Coyle, who has personally been active in pushing for updates to university economics curricula, suggests a 'puzzle to me is that a lot of the behavioural insights (and other shifts away from neoclassical Econ) is by now pretty mainstream in the academic discipline, so why is the policy world stuck with the previous paradigm?
5. For example, Monzo now enables customers to have up to 20 'jars'.
6. This raged as a major debate within behavioural economics, and early evaluations of the US 401k program (Madrian et al). Despite some early evidence suggesting that such a substitution might be occurring. Later work concluded that savings substitution was relatively limited – ie, mental accounts were alive and well.
7. William Nordhaus. Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not.
8. Under the Duty, firms should be open and honest, avoid causing foreseeable harm, and support you in pursuing your financial goals. You should expect:
9. Laibson, D. (1997). *Golden Eggs and Hyperbolic Discounting*. The Quarterly Journal of Economics, 112(2), 443-477.
10. Thaler, R. (1980). *Toward a Positive Theory of Consumer Choice*. Journal of Economic Behavior & Organization, 1(1), 39-60.
11. Tversky, A., & Kahneman, D. (1981). *The Framing of Decisions and the Psychology of Choice*. Science, 211(4481), 453-458.
12. Huber, J., Payne, J. W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. Journal of Consumer Research, 9(1), 90-98.
13. Lichtenstein, S., & Slovic, P. (1971). Reversals of Preference Between Bids and Choices in Gambling Decisions. Journal of Experimental Psychology, 89(1), 46-55.
14. These effects are often contingent on other beliefs, too, such as how Rutchick (2010) found the effects of churches on voting were only found among Christians. Researchers have also found that many of these effects appear to operate below conscious awareness.
15. For example, if you calculate the hours that people spend looking after their children or caring for the elderly, relatively, and value that at minimum wage levels.
16. March 13th, 2024