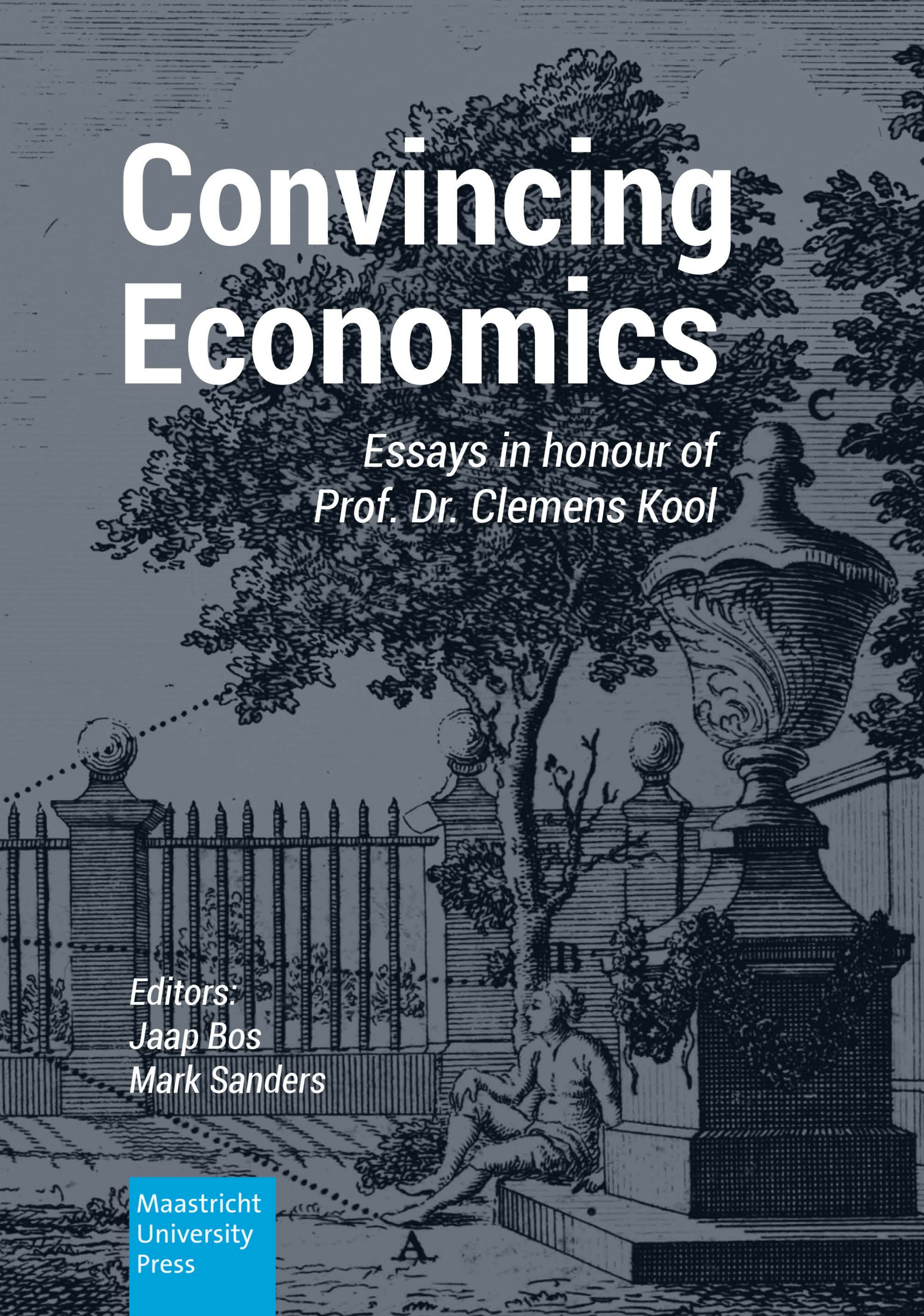


Convincing Economics

Essays in honour of Prof. Dr. Clemens Kool

Editors:
Jaap Bos
Mark Sanders

Maastricht University Press

Convincing Economics

Essays in honour of Prof. Dr. Clemens Kool

Convincing Economics

Essays in honour of Prof. Dr. Clemens Kool

editors

Jaap Bos
Mark Sanders

Convincing Economics
Essays in honour of Prof. Dr. Clemens Kool

Published by:
Maastricht University Press | Maastricht University, The Netherlands

Version: 2025-01
ISBN: 9789403769615
DOI: 10.26481/mup.2501

Download, read or buy the book here: umlib.nl/mup.2501

Cover Image:
"La chambre noire" (The Camera Obscura), originally published in The Universal Magazine of Knowledge and Pleasure, Vol. 10, 1752, p. 214.
Image source: Wikimedia Commons
This image is in the public domain.

Table of contents

Introduction:
In Search of Resilience – A Tribute to Clemens Kool

This volume is offered in honor of Clemens Kool, whose work has shaped monetary economics in the Netherlands and far beyond. As colleagues, collaborators, and students, we have come together to reflect not only on the topics Clemens has helped illuminate, but also on the spirit of intellectual curiosity, critical engagement, and methodological rigor that he has consistently exemplified throughout his career.

The essays in this liber amicorum span a range of subjects that Clemens has contributed to or engaged with over the years. While the thematic diversity is striking—from monetary policy and interest rate dynamics to convergence in the Eurozone, from inflation mechanisms to behavioral models under uncertainty—there are common threads running through the volume: a commitment to empirical grounding, a concern with robustness over elegance, and a recognition that in the face of complex and often unpredictable dynamics, economic systems must be designed to endure as much as to optimize.

In the opening chapter, Menno Middeldorp examines the predictive accuracy of the Federal Open Market Committee's projections. His analysis reminds us that while central bank transparency has increased significantly, the effectiveness of forward guidance remains mixed. Forecasting remains a difficult endeavor, particularly in uncertain environments—a theme that echoes throughout this book.

Harald Benink, Sara Murawski, and Mark Sanders explore the challenge of economic convergence within the Eurozone in Chapter 2. They argue that in the absence of adequate stabilizing mechanisms, divergence may persist or even intensify. Their call for a European investment agenda aligns closely with Clemens' long-standing interest in institutional design and macroeconomic policy coordination.

In Chapter 3, Casper de Vries reflects on the role of priors in economic modeling and connects Bayesian thinking across econometrics and game theory. His essay speaks to Clemens' appreciation for methodological clarity and the thoughtful application of statistical reasoning in economics.

Chapter 4, by Bikker, van Leuvensteijn, and Meringa, turns to the Dutch pension sector—a field where Clemens has provided expert insight. The authors analyze competition dynamics and efficiency, highlighting the limitations of structural reforms in sectors with restricted market forces. Their work underscores the importance of institutions in shaping economic outcomes—an area Clemens has examined extensively.

Peter Schotman revisits themes from his longstanding academic dialogue with Clemens in Chapter 5, re-evaluating structural breaks in short-term interest rates using modern econometric tools. This contribution is both a scholarly update and a recognition of the analytical foresight Clemens demonstrated in his early research.

In Chapter 6, Candelon and Hadzi-Vaskov analyze sovereign spreads in emerging markets, particularly in light of the COVID-19 pandemic. Their work illustrates how global shocks can alter convergence patterns and highlights the importance of understanding regional dynamics—an analytical lens that Clemens has consistently advocated in European monetary analysis.

The theme of uncertainty is further developed by De Grauwe and Ji in Chapter 7. Their model introduces radical uncertainty into macroeconomic analysis, showing how non-rational expectations and endogenous shifts in sentiment can amplify macroeconomic instability. The implications for monetary policy are profound: resilience may matter more than optimization—an idea that resonates with Clemens' emphasis on robust policy frameworks.

In Chapter 8, Meijers, Muysken, and van Veen explore inflation from a post-Keynesian perspective. Using a stock-flow consistent model, they analyze the role of import prices and mark-ups in shaping the distributional consequences of inflation. Their work reflects the integration of theoretical insight with empirical realism, a hallmark of Clemens' own approach.

Chapter 9, by Piccillo, returns to forecasting, applying a Markov switching model to study exchange rates and risk aversion. The analysis draws connections with earlier chapters in highlighting the limits of predictive models under changing regimes—a recurring theme in this volume.

Finally, Bos, van Lamoen, and Yuan (Chapter 10) examine how micro-level data can enhance macroeconomic forecasting. Their study demonstrates how firm-level heterogeneity, when properly modeled, can improve aggregate predictions. The chapter highlights the potential of new empirical tools in advancing the field—a perspective Clemens has often championed.

Taken together, the contributions in this volume reflect a shared concern: that the search for ever more accurate models may be less fruitful than building economic systems that are robust to the unknown. The notion of resilience—structural, institutional, and conceptual—emerges as a unifying thread. Rather than seeking to eliminate uncertainty, the task may be to design frameworks that can withstand it.

Clemens Kool has consistently brought this kind of balanced perspective to economic research and policy discussion. With a deep appreciation for both theoretical elegance and empirical rigor, he has helped shape how economists think about monetary systems, macroeconomic stability, and institutional design. His work and mentorship have left a lasting impact on several generations of economists.

As Clemens transitions to the next chapter of his life, we extend our deepest thanks and appreciation. This volume is a testament to his scholarship, but also to the collegiality, generosity, and intellectual seriousness that have defined his career. We hope it offers not only reflection, but inspiration.

How hot is the dot plot?
The accuracy and informativeness of FOMC projections

Menno Middeldorp

Introduction

Along with other central banks around the globe, the US Federal Reserve (Fed) became increasingly transparent and communicative about its policies and forecasts from the 1990s to the 2010s. Fed policy makers and other central bankers believed that sharing information would have economic benefits. See Middeldorp (2010) for a historical analysis of the Fed's motivations for becoming more transparent. The academic literature shows empirical support for this belief.

Regarding the financial markets in particular, communication is seen as improving the predictability of monetary policy, thereby lowering market volatility, and contributing indirectly to a more stable economy. To this end, many central banks communicate about future policy rates directly, often called "forward guidance." See Moessner, Jansen and de Haan (2017) for a survey of forward guidance in theory and practice. Often this includes publishing projections of policy rates, which several central banks including the Fed do.

The exact nature of the projections released by central banks differ. The Fed's policy committee, the Federal Open Market Committee (FOMC), meets to set interest rates every six weeks, followed by a statement and a press conference. Once every quarter the Fed also releases an anonymous survey of the economic and interest rate projections of the FOMC participants in the Summary of Economic Projections. This is illustrated with a chart that shows one unnamed dot for each participant for their assessment of the "appropriate" fed funds rate at the end of the current and next two to three calendar years, which often called the "dot plot." The FOMC participants are not actually asked to predict the future policy rate (that would require them to not only to make an assessment of appropriate policy but also of what FOMC members are likely to vote) but the projections as a group can be seen as a prediction assuming that members vote in line with their assessment of appropriate policy.

Forward guidance has been found to impact financial markets, see Brubakk, ter Ellen and Xu (2021) for a recent analysis and references to earlier studies. Bongard (2021) and Couture (2021) show that markets responded to the "dot plot", in the early 2010s. Whether forward guidance also helps financial markets to predict monetary policy is less clear. Kool and Thornton (2015) find that guidance does help for some central banks but only for short term rates and not for the Fed. Some like Morris and Shin (2002,2018), Kool, Middeldorp and Rosenkranz (2011) and Ehrmann et al (2019) raise concerns that signals about future policy from a central bank could distract or crowd-out private information of market participants and lead to markets that are actually less able to predict monetary policy, although the evidence is mostly theoretical.

To assist markets in better predicting policy rates it could help if central bankers are good at forecasting themselves. This is especially important if the central bank signal does indeed crowd-out private information, because – in line with Kool, Middeldorp, Rosenkranz (2011) and Ehmann et al (2019) – it can then only improve the predictive ability of the market if it provides better information than what it displaces.

This paper thus asks how the accuracy of the FOMC dot plot compares to that of private sector forecasters and the market. It also examines whether market or private sector forecasters become better at predicting the policy rate after seeing the dot plot.

Data

The Fed releases its projections every quarter. The dot plot shows individual assessments of the fed funds rate range midpoint that the FOMC participant expects will be appropriate for the end of the current calendar year, the subsequent year, and the year after. The plot also displays assessments for the "longer run" and (after FOMC meetings in the latter part of the year) for the third year in the future; these are not included in the analysis below because private sector forecasts are not available that far ahead. Bloomberg provides historical data for the median and average projection of members.

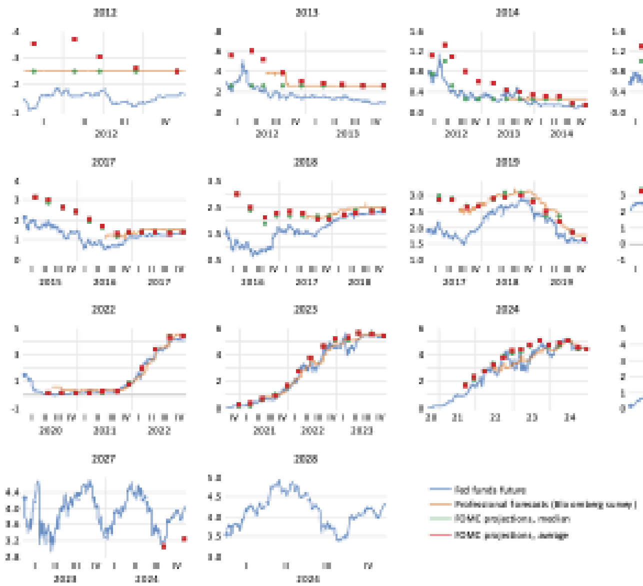
There are also many private sector economists and market strategists who provide forecasts of the fed funds rate target range ceiling to a survey by Bloomberg. The survey is more frequent than that of the FOMC and forecasters can notify Bloomberg of updates at any time. Bloomberg also provides historical averages of the private-sector forecasts for the upper bound of the FOMC target range for every day for the current, next, and subsequent year. The Bloomberg average gives more weight to recent forecasts to reduce the impact of stale contributions.

Fed funds futures are a common measure of the market's expectation of future monetary policy. They are traded contracts on the average fed funds rate for a specific upcoming month. That is different from the FOMC projections or private-sector forecasts because these refer to respectively the fed funds target range midpoint and ceiling. The Fed uses its operations to steer the fed funds rate near to the target. The fed funds future rate reflects more than just market expectations but also compensation for risk, so there is likely a term-premium in the fed funds future rate. Diercks and Carl (2019) and Kim and Tanaka (2016) show that estimates of the term premium not only vary widely across models and methods but also across time, although they are probably close to zero up to a few months out.

In the interests of brevity this paper refers to the FOMC projections, private sector forecasts and fed funds collectively as "predictions".

Figure 1 shows how the predictions for a particular year-end developed over time. The time-series converge towards the actual target rate, but there are still minor differences between the predictions because of the slightly different predicted rates, as noted above. Over this sample there is a consistent difference of 12.5bp between the midpoint and the ceiling. The difference between the effective rate and the midpoint fluctuates but, on average, is 2bp.

Figure 1: Historical FOMC projections, private sector forecasts and fed funds future rate per predicted date[1]



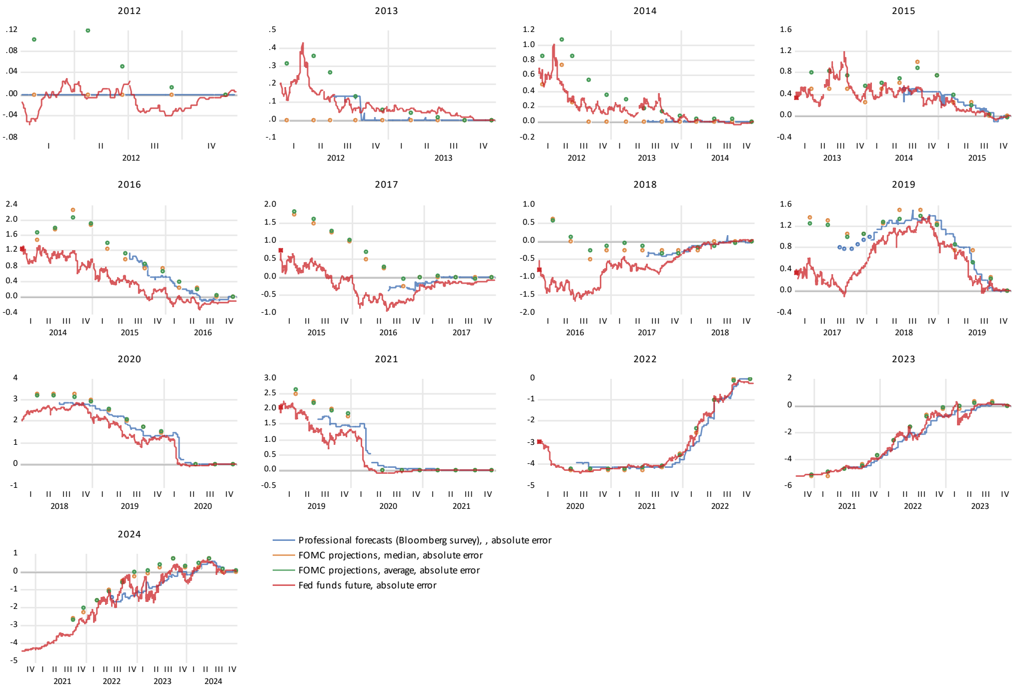
Source: Bloomberg, Macrobond

[1] BB = Bloomberg private sector forecast weighted average for the fed funds target range ceiling
FF = Fed funds future rate for the average effective fed funds of the last month of the year
FedMed = FOMC median projection of appropriate fed funds target range midpoint
FedAv = FOMC average projection of appropriate fed funds target range midpoint
Source: Bloomberg, Macrobond

To see how accurate the above predictions are, errors are calculated for each relative to the matching realization at the end of each calendar year:

FOMC projection average/median – actual target range midpoint at the end of the predicted year
Bloomberg private sector average – actual target range ceiling at the end of the predicted year
Fed funds future rate – average effective fed funds rate last month of the predicted year

Figure 2: Historical accuracy of predictions per predicted year-end fed funds rate[2]



How accurate are the FOMC projections relative to other predictions?

To determine the accuracy of the FOMC projections relative to the other predictions, this section examines the magnitude of the prediction errors as the predicted year-end date approaches. Four panel regressions are used. The dependent variables are the absolute values of the difference in the prediction error between the two variants of the FOMC projections (median and average) versus either the Bloomberg forecasts or the fed funds futures rates. There are four different regressions as a result. In all of them the independent variables are dummies to indicate which calendar year the predictions are in relative to the calendar year which is predicted (y). The periods in the panel are all weekdays (t) on which the FOMC releases projections (set P). The cross-sections are the predicted year-ends (y). For the Bloomberg forecast average and fed funds rate the data used are for the previous day (t-1) to avoid them reflecting information from the latest FOMC projections.

There are four dependent variables

1	$\lvert a_{t,y}\rvert - \lvert b_{t-1,y}\rvert$	3	$\lvert m_{t,y}\rvert - \lvert b_{t-1,y}\rvert$
2	$\lvert a_{t,y}\rvert - \lvert f_{t-1,y}\rvert$	4	$\lvert m_{t,y}\rvert - \lvert f_{t-1,y}\rvert$

That are regressed on these independent variables.

$$\ldots = \alpha d_{t,y}^{y-2} + \beta d_{t,y}^{y-1} + \gamma d_{t,y}^{y} + \epsilon_{t,y} \ \text{ if } t \in P$$

Where:

y	calendar year of projection
t	trading day
P	set of days when projections are released

$|m_{t,y,}|$ absolute value of the median FOMC projection error (m) for the end of calendar year if the trading day is a member of the set of days when projections are released

$|a_{t,y}|$ same for average FOMC projection error (a)

$|b_{t-1,y}|$ absolute value of the Bloomberg forecast error (b) of the day before the FOMC projections

$|f_{t-1,y}|$ absolute value of the fed funds future rate (f) of the day before the FOMC projections

$d_{t,y}^{y+0,1,2}$ dummies for if the predicted year-end day y is the same year as day t $(d_{t,y}^{y})$ the next year $(d_{t,y}^{y-1})$ or the year after $(d_{t,y}^{y-2})$

The coefficients of the dummies show the difference in accuracy between the FOMC projections versus professional forecasters and the market. If these are negative it means that the FOMC is more accurate. Statistical significance is tested using standard errors that are robust for serial correlation, cross-section correlation and heteroskedasticity.

Table 1: Panel regressions results

		Predicted year dummies			Periods	Cross-sections	Observations
Year of data relative to predicted year-end (y)		y-2	y-1	y			
Prediction horizon (quarters)		8 - 12q	4 - 8q	<4q			
Difference in errors							
FOMC							
Average	Bloomberg	-0.11	-0.06	-0.01	52	13	97
Median		-0.09	-0.06	-0.02			
Average	Futures	0.29	0.05	0.00	52	13	144
Median		0.23	0.01	-0.02			

The FOMC is more accurate than the professional forecasts but less accurate than the market, especially at longer horizons, but none of these differences are statistically significant. It is an open question whether adjusting for term premiums would result in the futures significantly outperforming the FOMC, although that would require additional assumptions and modelling.

Do private sector forecasters and the fed funds futures market find the FOMC projections informative?

This section analyses if the FOMC projections help private sector forecasters and the market to better predict the fed funds. The change in the absolute prediction error around FOMC dates is calculated. If it is significantly different from the typical decline in the error across all days, then the information provided by the FOMC meeting are deemed to improve the predictability of the policy rate. Here too the improvement in accuracy as the prediction horizon declines is accounted for using dummies.

The fed funds future market should be able to adjust quickly to any news, so the change in the fed funds future rate over one day is used. However, forecasters may not immediately adjust and submit their forecasts to Bloomberg, so here the change over three days is used.

The dependent variables are now:

$|b_{t+3,y}| - |b_{t-1,y}|$ change in accuracy of the Bloomberg forecasts from $t-1$ to $t+3$

$|f_{t,y}| - |f_{t-1,y}|$ change in accuracy of the fed funds future rate

While the independent variables are the following:

$$\ldots = (\alpha d_{t,y}^{y-2} + \beta d_{t,y}^{y-1} + \gamma d_{t,y}^{y}) + (\alpha_p p_t d_{t,y}^{y-2} + \beta_p p_t d_{t,y}^{y-1} + \gamma_p p_t d_{t,y}^{y}) + \epsilon_{t,y}$$

Where:

p_t dummy for the days *t* that projections are released

The coefficients in the first part (α, β, γ) are the average change in the prediction errors for all days within the three horizons. Interacting the dummies for the prediction horizon and the FOMC projections means that the coefficients in the second part (α_p, β_p, γ_p) of the regression can be seen as the *additional* change in the prediction error around the release of projections. This makes it possible to provide an answer to the question whether the release of the FOMC projections improves the accuracy of the Bloomberg average and the fed funds future more than the typical improvement in accuracy on other days. Although, the answer is not definitive because the FOMC projections are not the only information released by the FOMC.

Table 2: Regression results

		y-2		y-1		y		cross-sections	periods	observations
									(with projections)	
Bloomberg (t+3)-(t-1)	all days	-0.006		-0.014	**	-0.011	**	13	(52)	3158
	+projections	-0.024		-0.009		-0.010			(97)	3372
FF (t)-(t-1)	all days	-0.001		-0.002	**	-0.002	**	13	(52)	6242
	+projections	-0.012		0.006		-0.002			(144)	8756
*p<0.1 **p<0.05 ***p<0.01 White two-way cluster covariance										

The results show that as the predicted year-end approaches the pace at which accuracy improves increases. That reflects the convergence of the errors to close to zero in Figure 2.

The difference between the "all days" coefficients for the forecasters and market reflects that the former is over four days (t+3 vs. t-1) while that for the fed funds future is just one day. Also, the samples are also not identical.

The coefficient for the projections days shows the *additional* change in accuracy. So, in all cases the accuracy improves faster than on other days. The improvement is greatest for the days in years furthest away from the prediction date. However, none of these additional improvements around the release of the FOMC projections are statistically significant.

Interpretation and discussion

The FOMC projections appear to be more accurate than private sector forecasters but not the fed funds rate, but this cannot be demonstrated with statistical significance. The projections also appear to be informative for forecasters and the market, especially at longer horizons, but here too the results are not statistically significant.

The lack of statistical significance is likely related to the limitations of the sample. There are only fifty-two releases of projections. More importantly, there are only thirteen years in which the year-end fed funds rates are actually realized. Using a panel with the predicted years as cross-sections yields more observations, but the errors are also correlated across the cross-sections and within the time-series. These effects are accounted for by robust standard errors. So, the subsequent low significance appropriately reflects the lack of independence of the limited observations available.

Putting statistical significance aside, the result that the FOMC appears better at predicting itself than professional forecasters is not surprising (Table 1). Even if the Fed was not better informed or more skilled at forecasting the economy the FOMC obviously has more insight into how it is likely to translate its economic outlook into monetary policy. For the same reason it is plausible that forecasters and the markets find the FOMC informative (Table 2).

For the current year, the median FOMC projection turns out to be a slightly better predictor of the actual future rate than the average FOMC projection. That could reflect that there is often a consensus in the committee on appropriate policy in the short term. A few dissenting views are reflected in the average, but they do not change the outcome, which is better reflected by the median view of the participants. Note also that the projections show the assessment of *appropriate* policy. They are not the forecast of the FOMC participant on what the actual outcome will be.

Although the FOMC projections are the best available predictor at short horizons, the forecasters and markets appear to find the FOMC to be most informative at longer horizons. One reason may be that there is a lot of information about the near-term outlook for monetary policy provided by FOMC participants through other means. At every FOMC meeting there is a statement and a press conference. Minutes are released after the meetings. FOMC participants give speeches, Congressional testimony, and media interviews. The focus of all of these tends to be on near-term economic conditions and policy outlook. The FOMC projections are one of the few moments when FOMC participants share information about the longer-term economic outlook and policy rates.

The market appears to be the best predictor of the fed funds rate at longer horizons and the FOMC (median) does not significantly outperform it at any horizon (Table 1). The likelihood that a term-premium biases the future rate means that it is possible that the fund funds market is even better at predicting the fed funds rate than the (insignificant) results suggest. From the perspective of mainstream financial literature, the idea that this type of market is the most accurate prediction available is not a surprise. Market participants have strong incentives to strive to accurately predict the future underlying rate given they have money on the line and bear the losses and reap the benefits of their prediction accordingly. The market price also incorporates both the public information available, and the private information and assessments of many participants, as discussed.

Conclusion

The analysis yields the plausible results that the FOMC is better at predicting its own future monetary policy than professional forecasters. The market, however, typically outperforms the FOMC, except at the shortest horizon. Forecasters and the markets both appear to find the FOMC projections informative. However, none of these results are statistically significant reflecting the limited number of correlated observations in the sample.

References

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., & Jansen, D.-J. (2008). *Communication about future policy rates in theory and practice* [Working paper]. University of Groningen.

Bongard, M., Galati, G., Moessner, R., & Nelson, W. (2021). Connecting the dots: Market reactions to forecasts of policy rates and forward guidance provided by the Fed. *International Journal of Finance & Economics, 26*(4), 684–706.

Brubakk, L., ter Ellen, S., & Xu, H. (2021). Central bank communication through interest rate projections. *Journal of Banking & Finance, 124*, Article 106044.

Couture, C. (2021). Financial market effects of FOMC projections. *Journal of Macroeconomics, 67*, Article 103279.

Diercks, A., & Carl, U. (2019, January 8). A simple macro-finance measure of risk premia in fed funds futures. *FEDS Notes*.

Ehrmann, M., Gaballo, G., Hoffmann, P., & Strasser, G. (2019). Can more public information raise uncertainty? The international evidence on forward guidance. *Economics Letters, 181*, 145–149.

Federal Reserve. (2024). *Summary of economic projections, December 18, 2024*.

Kim, D. H., & Tanaka, H. (2016, November 18). Front-end term premiums in federal funds futures rates and implied probabilities of future rate hikes. *FEDS Notes*.

Kool, C. J. M., & Thornton, D. L. (2015). How effective is central bank forward guidance? *Federal Reserve Bank of St. Louis Review, 97*(4), 303–322.

Kool, C., Middeldorp, M., & Rosenkranz, S. (2011). Central bank transparency and the crowding out of private information in financial markets. *Journal of Money, Credit and Banking, 43*(4), 767–799.

Middeldorp, M. H. (2010). *Central bank transparency, private information and the predictability of monetary policy in the financial markets: Theoretical, experimental and empirical evidence* (Doctoral dissertation). Utrecht University.

Morris, S., & Shin, H. S. (2002). Social value of public information. *American Economic Review, 92*(5), 1521–1534.

Morris, S., & Shin, H. S. (2018). Central bank forward guidance and the signal value of market prices. *AEA Papers and Proceedings, 108*, 572–577.

Coming Together or Drifting Apart:
The Case for a European Investment Agenda.

Harald Benink, Sara Murawski and Mark Sanders

> *"But it would be much better to tackle the root cause of this growth divergence. These differences between North and South are not after all a God-given natural phenomenon."*
>
> *— Klaas Knot, 2020*

Introduction

In 1976, Lucas formulated his famous Critique that it is naïve to try to predict the effects of a change in economic policy entirely based on relationships observed in historical data. He basically argued that, as policies change, so will these relationships. In close analogy the Clemens' Critique can be stated as, it is naïve to assume your analysis will convince Clemens Kool on the first read. The inevitable "I am not convinced" is the phrase that implies you will need to do a better job of making your case. Therefore, in honor of Clemens Kool, we decided to take a report that was prepared by Sara and commented by Clemens, attempting to convince him of the arguments in a second try. We leave it to the reader (Clemens, first and foremost) to pass judgement on our success.

This chapter is about convergence in the European Union and more specifically the European Monetary Union (EMU), established with the Treaty of Maastricht. Convergence had a central role in the establishment of the EMU. The 1989 Delors-report argued that a monetary union requires a sufficient degree of convergence of economic performance and (hence) economic policies. It states that "[p]arallel advancement in economic and monetary integration would be indispensable in order to avoid imbalances which could cause economic strains and loss of political support for developing the Community further into an economic and monetary union."[2] The Maastricht Treaty (1992) is even more explicit, arguing that "[t]he Community shall have as its task...a high degree of convergence of economic performance."[3] But what kind of convergence is meant? And which mechanisms in the design of the EMU were thought, hoped or believed to deliver such convergence?

Convergence

The convergence debate can be traced to the mainstream, neoclassical growth model. In this model, Solow (1956) assumed diminishing returns to capital, linear depreciation and a constant saving rate and showed that these ingredients would cause an economy to converge to a steady state level of capital and therefore income per capita. This convergence is conditional on parameters and exogenous variables such as population growth, saving rate, depreciation rate and productivity levels and growth. Provided these are comparable across economies, the model predicts they will end up in comparable steady state growth paths. The empirical performance of this model in global cross-sectional data is impressive (Mankiw, Romer and Weil, 1994) and convergence, especially among similar and integrated economies, is quite strong (e.g. Barro and Sala-I-Martin, 2004).

Opening the economies to trade and free flows of capital and labor, was therefore believed to speed up the convergence process by ensuring knowledge flows freely and the tide of productivity growth could lift all boats. Giving up national currencies to eliminate exchange rate fluctuations and uncertainty would facilitate trade and investment across EU borders and thereby push the poorer parts of the Union onto higher and steeper growth trajectories to rapidly catch up with the core. Transitional growth explains how Germany and Japan rapidly recovered from World War II and why countries like the Asian Tigers and Dragons rapidly converged to the

---

2 https://ec.europa.eu/economy_finance/publications/pages/publication6161_en.pdf
3 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A11992M%2FTXT

industrialized world after adopting the "right" institutions and integrating into the global (Western) economy.

The Solow model fails to explain two important facts. Why would some countries/economies converge to a slower growth than others and why would economies fail to converge to the global frontier? In the context of the European Union, why do some regions in Europe lag the levels and growth rates of GDP per capita in others, what causes some to suffer recession while others enjoy a boom and why are some affected more and others less by similar external shocks.

Before we turn to the empirical evidence and experiences in Europe over decades since he Maastricht Treaty, we briefly review a textbook endogenous growth model in the spirit of Romer (1990) with an extension due to Jones and Vollrath (2024). This helps us understand the mechanisms that may cause con- and divergence in the level and growth of GDP per capita in the European Union as well as the transmission of positive and negative shocks to (parts of) the Union.

A toy model of growth and convergence

If we assume a standard Cobb-Douglas production function:

$$Y_i = K_i^{\alpha}(A_i h_i L_i)^{1-\alpha} \quad (1)$$

Where Y, K, A, h and L represent output, capital, labor augmenting productivity, human capital and labor respectively, α is the output elasticity of capital between zero and one, i indexes the country and we suppress time subscripts (and assume human capital is constant). Solow (1956) added an exogenous constant population and productivity growth and the standard capital accumulation function:

$$A_i = A_{i0}e^{g_{Ai}t} \quad (2)$$

$$L_i = L_{i0}e^{g_{Li}t} \quad (3)$$

$$dK_i = s_i Y_i - \delta K_i \quad (4)$$

Where t is time and 0 indexes the exogenous starting values, s is the gross fixed capital formation rate and δ is the depreciation rate. We know (e.g. Jones and Vollrath (2024, ch. 2)) that the steady state growth in output per capita is given by:

$$g_{Y/L} = g_{Ai} \qquad (5)$$

And output per capita on the balanced growth path will converge to:

$$y_i^{BGP} = \left(\frac{s_i}{\delta + g_{Ai} + g_{Li}}\right)^{\frac{\alpha}{1-\alpha}} h_i A_{i0} e^{g_{Ai} t} \qquad (6)$$

From this simple model we can already see that convergence between two economies i and j in terms of growth rate and level of GDP per capita, will depend on the values of the parameters and exogenous variables in the model. But this model assumes closed economies, such that convergence in per capita GDP growth rates or levels would be a pure coincidence. All economies i, however, do converge to their own steady state and balanced growth path, such that if we assume gAi to be equal to or converge among these economies, we predict convergence in growth rates and by (6) convergence to a constant relative output per capita. Moreover, within the European Union, markets, notably goods, labor and capital markets, have all been integrated. To the extent that labor and capital are mobile between economies i and j, this would imply convergence in wages and cost of capital, such that, assuming competition in factor markets, also the marginal products of capital and labor would converge. Endogenizing gross fixed capital formation rates would then imply higher levels of investment in the capital poor economies, while labor mobility would imply that all economies effectively face the same overall growth rate in workforce (as people will move wherever their marginal productivity is highest).

Of course, labor is far from perfectly mobile across the Union and risk premia still distort capital markets. We can therefore derive more useful insights from the model in Jones and Vollrath (2024, ch. 5 and 7) that assumes labor and capital are immobile, but knowledge accumulates endogenously (Romer, 1990) and knowledge spillovers exist between a frontier and a lagging economy.

The model above is first extended by replacing (2) by:

$$dA_i = \theta L_{Ri}^{\lambda} A_i^{\varphi} \quad (7)$$

Where (7) describes the process of knowledge accumulation at the frontier country i as a function of the labor allocated to innovation, LRi, and the stock of knowledge accumulated in the past, Ai and θ, λ, φ are positive parameters and the last two are less than one. For the lagging economy j we assume:

$$dA_j = \psi h A_i^{\gamma} A_j^{1-\gamma} \quad (8)$$

Where it should be noted that $A_j \leq A_i$ must hold and productivity in the lagging country benefits from innovations at the frontier, technology adoption in the past and the level of human capital, h in the labor force. Jones and Vollrath (2024) shows that on the balanced growth path we have:

$$g_A^{SS} = g_{Aj} = g_{Ai} = \frac{\lambda}{1-\varphi} g_{Li} \quad (9)$$

Whereas the outputs per capita will converge to:

$$y_i^{BGP} = \left(\frac{s_i}{\delta + g_A^{SS} + g_{Li}}\right)^{\frac{\alpha}{1-\alpha}} (1 - s_R)\, h_i A_{i0} e^{g_A^{SS} t} \quad (10)$$

$$y_j^{BGP} = \left(\frac{s_j}{\delta + g_A^{SS} + g_{Lj}}\right)^{\frac{\alpha}{1-\alpha}} h_j \left(\frac{\psi h_j}{g_A^{SS}}\right)^{\frac{1}{\gamma}} A_{i0} e^{g_A^{SS} t} \quad (11)$$

Such that the model no longer predicts convergence in GDP per capita levels, even if all parameters and exogenous growth rates are the same or somehow would converge to the same levels due to market integration. The model does predict convergence in growth rates and all else equal, poorer economies that start out further below their BGP will grow faster. But in this model, it is perfectly possible for economies to show no convergence in output per capita or even experience episodes of divergence following changes in the structural parameters that permanently shift and/or shock economies off their balanced growth paths. Since shocks and structural change have been a constant in the European Union since before the Treaty of Maastricht, convergence between parts of the Union should not be expected without deliberate policies to achieve it, even if full market integration is achieved.

For such policies, however, the Lucas (and Clemens?) Critique applies. In implementing European policies, the fundamental parameters and possibly relationships among the exogenous variables would change. Consequently, the behavior of the system we are trying to influence, will change as well. Policy makers should therefore not design and execute but should instead experiment and adapt. It certainly is not enough to force nominal variables (wages, prices, inflation, interest rates, debt and deficit ratios) to converge and/or remain converged to ensure real convergence. And even if real convergence is achieved in the long run, the European Union would still need powerful, automatic stabilizers and risk sharing mechanisms to ensure that shocks dissipate and business cycles remain converged across the Union.

In the below, we will first discuss three types of convergence: nominal, real and cyclical convergence.[4] Then we turn to some modest policy interventions that we believe will bring European economies closer together and empower Europe/the EU to address the urgent challenges the continent faces.

[4] Following IMF Working Paper (2018). Economic Convergence in the Euro Area: Coming Together or Drifting Apart?

Nominal, real and cyclical convergence

Nominal convergence refers to convergence in nominal variables. In case of the EMU, the most famous are the Maastricht convergence criteria, used to determine if a country is ready to introduce the euro. The Maastricht convergence criteria require four things: price stability, sound and sustainable public finances (consisting of both a deficit and debt criterion), exchange-rate stability and low and stable long-term interest rates.[5] The criteria were introduced in the run-up to the euro, to achieve price stability and enforce converging inflation across the EU.

Real convergence refers to the convergence of real variables, notably income levels between countries. Real convergence refers to the process where countries with low-income levels (measured by GDP per capita) "catch up" with countries with high income levels. Two types of income convergence can be distinguished: beta-convergence, which measures the degree to which countries with lower GDP per capita levels grow faster than countries with higher GDP per capita levels; and sigma-convergence, which is the decline in the dispersion of GDP per capita levels between countries over time.

Finally, we can distinguish cyclical convergence, measuring the correlation in real variables over the business cycle across European Member States to see if booms and busts are of similar size and timing across the Union.

The Maastricht Treaty

Convergence and prosperity were enshrined as the objectives of the EU in the Maastricht Treaty. Both are strongly related to real convergence (Creel, 2018). Significantly, however, real convergence targets were absent in the Maastricht treaty. Instead, real convergence was assumed to follow from nominal convergence, market discipline and further integration. Additionally, the nominal Maastricht convergence

[5] https://www.consilium.europa.eu/en/policies/joining-the-euro-area/convergence-criteria/

criteria were meant to reconcile the different fiscal policies across the EMU with a common monetary policy (IMF, 2018). Nominal convergence, however, does not automatically lead to real convergence.

On the contrary: a narrow focus on nominal convergence can hamper real convergence, for example if it leads to contractionary fiscal policy that negatively impacts the weaker economies. Paul de Grauwe argued in 1996 that complying to the Maastricht convergence criteria before entering the euro would maximize the costs of convergence and could actually lead to increasing inflation and higher deficits (de Grauwe, 1996).[6] In 1997, a group of 70 Dutch economists pointed out that the introduction of the euro forced countries to implement austerity policies with detrimental socio-economic effects.[7] In Italy, for example, fiscal consolidation in the 1990s and 2000s was much stronger than in other advanced economies (IMF, 2011) and this went hand in hand with a deteriorating standard of living in terms of GDP per capita since the 2000s (Beun, 2022). But because nominal convergence is easier to measure and observe, the Maastricht Treaty set nominal convergence criteria to determine who could join the euro in the first wave.

Nominal convergence in the euro zone

Before the introduction of the euro, inflation rates converged significantly between the twelve original euro members. The average inflation rate was 1.3% in 1998 compared to 3% in 1995 (IMF, 2018). Since then, the inflation convergence of the EU12 has stalled

---

[6] The reasoning is as following, taking the example of Italy: Because of its reputation, Italy will find it hard to convince the market that it will bring its inflation down sufficiently in due time before entering the euro. While bringing down inflation, unemployment rises. Meanwhile, the lira experiences real appreciation because its exchange rate is pegged to the euro. This causes more scepsis regarding the markets' disinflationary expectations and speculation is set in motion. Italy is forced to devalue the lira as a result, implying diverging inflation. As a result, nominal interest rates will remain high, and the debt burden increases.

[7] "Het huidige beleid is echter gefixeerd op de uitgavenkant: terugdringing van de collectieve uitgaven. Dat brengt omvangrijke bezuinigingen met zich mee die veel sociale en economische schade veroorzaken. Niet in de laatste plaats omdat veel lidstaten al sinds 1982 in een web van bezuinigingen gevangen zitten [...] De EMU-criteria en stabiliteitspactverordeningen werken de facto juist een Europees procyclisch beleid in de hand. Daardoor loopt met name de ontwikkeling van de werkgelegenheid in gevaar." https://www.volkskrant.nl/economie/met-deze-emu-kiest-europa-verkeerde-weg~bf1ca466/

(see Figures 1 and 2). The price levels of countries that joined the euro in 2007 or later (Eastern countries) converged only slightly.

Figure 1: Inflation Convergence (IMF 2018)

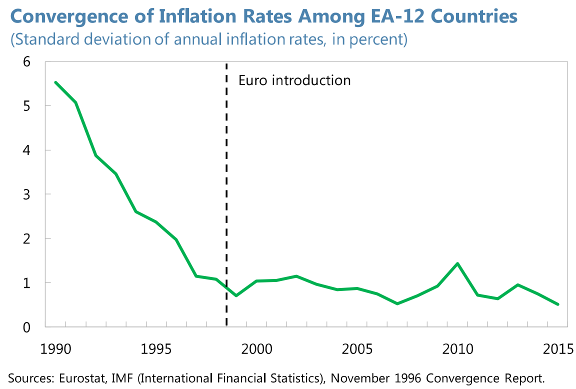
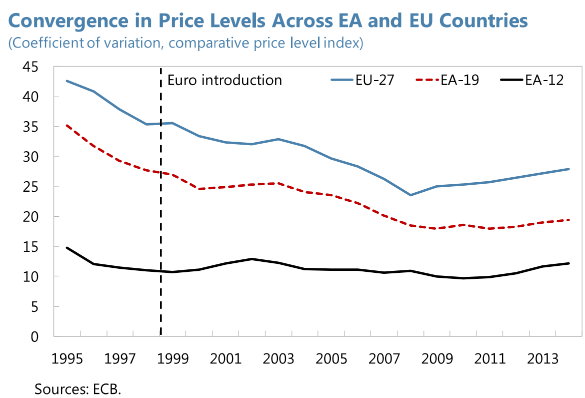
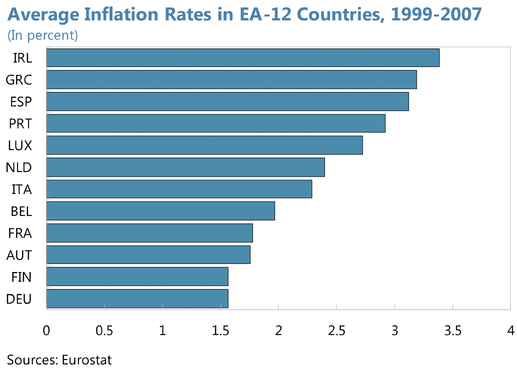


Figure 2: Price Convergence (IMF 2018)



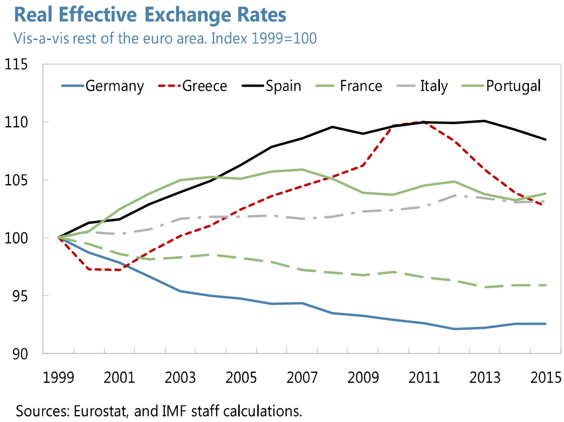
The (small, but persistent) differences between the inflation rates of the EA 12 countries left their marks. The relatively high inflation rates of (mainly) Southern countries negatively impacted their competitiveness, as visible in the development of their real effective exchange rates (IMF, 2018, see Figures 3 and 4). This is an important fact when reflecting on the question to what extent nominal convergence leads to real convergence. In this case, fulfilling the nominal convergence criteria for an important part did not prevent the real exchange rates to diverge.[8]

Figure 3: Inflation Rates (IMF 2018)



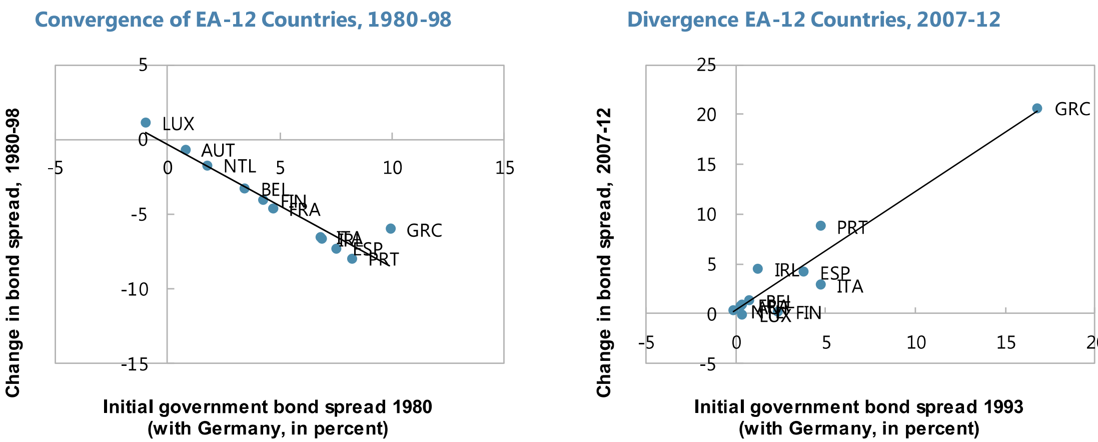
[8] The Clemens Critique: Het stuk over inflatieconvergentie is wel erg kort door de bocht. Vóór de start van de euro ging het erom om de transitie soepel te maken en alle NCBs te dwingen om een convergerend monetair beleid te voeren. Na de invoering speelt er iets heel anders en is (mijns inziens) de vraag in hoeveel jaren een land het inflatiecriterium niet haalt, niet relevant. Door het gezamenlijke monetaire beleid wordt de reële wisselkoers een belangrijk (het belangrijkste) instrument om evenwicht te realiseren bij asymmetrische schokken. Omdat de nominale wisselkoers is 1 is betekent dat dat je moet verwachten (en ook waarderen) dat er tijdelijke inflatieverschillen ontstaan die zorgen voor aanpassing. Voor landen met een gelijk ontwikkelingsniveau wil je wel dat over een langere tijd de gemiddelde inflatie ruwweg hetzelfde is. Tweede issue is dat verschillen in ontwikkeling (GDP per capita) leiden tot persistente inflatieverschillen. Landen die moeten inhalen (starten op een laag niveau) krijgen dan "automatisch" een hogere inflatie en reële appreciatie (in de literatuur bekend als het Balassa-Samuelson effect dat ontstaat omdat de prijs van niet verhandelbare goederen en diensten relatief stijgt als een land welvarender wordt). Ook dat kun je dus niet wegzetten als een (vermijdbaar) probleem. Tenslotte nog een punt over prijsniveaus: ook hier is de vraag over het gebrek aan volledige convergentie een issue is. Ook in Nederland (en andere landen), is het prijsniveau tussen regio's behoorlijk verschillend (denk met name aan onroerend goed, maar ook aan een "terrasje").
Our reply: Of course, we agree with Clemens that inflation convergence before and after the introduction of the euro is something different. Indeed, after the Euro, a sustained inflation differential implies a real exchange rate change that may indeed be necessary. The fact that these figures appear in Post-Euro Convergence reports by IMF and ECB suggest, however, that at least some had expected/hoped that the convergence in price levels and inflation would carry over beyond the introduction of the single currency such that the single currency would not create tensions. More importantly, it is important to note and show that it is the weaker countries in the Eurozone that experience an appreciation, increasing tensions and imbalances instead of adjusting to structural trends. If real exchange rate adjustments were needed to counterbalance persistent differences in GDP (growth), one would expect the faster growing and more advanced economies to appreciate. We have adjusted the text to prevent the reader from getting the impression that we consider this a preventable problem. We do believe, however, that the architects of the Eurozone did not anticipate the problem to take this form.

Figure 4: Real Effective Exchange Rates (IMF 2018)



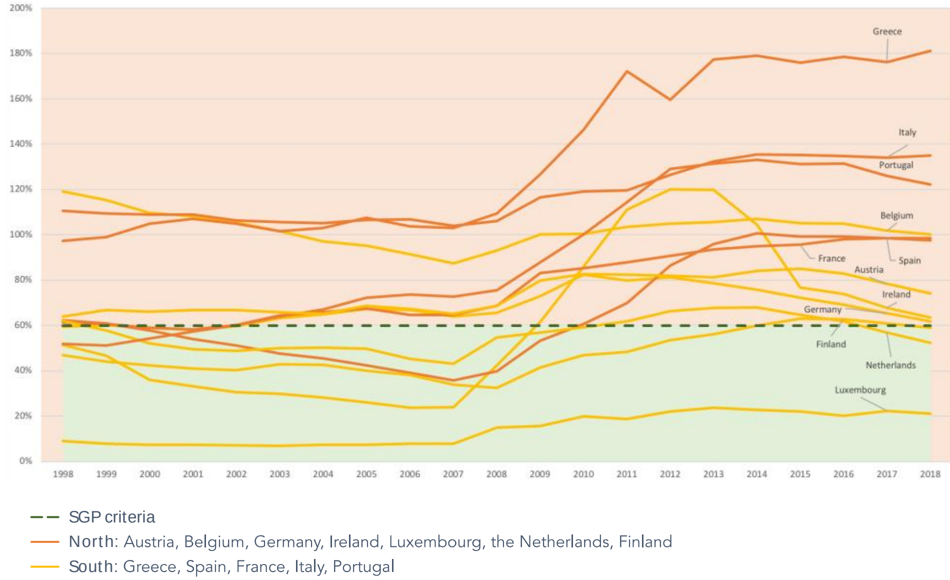
Nominal interest rates converged before the crisis and remained so in the first ten years of the euro, but this trend was reversed in the euro crisis (IMF, 2018; see figures 5).

Figure 5: Con- and Diverging Interest Rates (IMF 2018)



As for debt sustainability, quite some countries approximately met the 60% debt threshold when the euro was introduced (see Figure 6). However, all debt levels rose during the euro crisis, in most cases far beyond 60%, and divergence was and remains strong.[9]

Figure 6: Debt to GDP ratio (SFL 2020)



Debt continued to increase because the reference value of 3% of GDP for the deficit was breached 34 times between 1999 and 2007, including by France and Germany (Baldwin et al., 2015).[10]

---

[9] The Clemens Critique: Divergentie kan ontstaan door divergentie in (lange termijn verwachtingen van) inflatie en door divergentie in risicopremies. Mijn interpretatie van de literatuur is dat de divergentie in lange rentes voor het overgrote deel komt door risicopremies. Dat kun je niet oplossen met monetair beleid.
Our Reply: We agree with Clemens that this is the case. We made sure never to claim that this can be solved with monetary policy. Fact remains that in the Eurocrisis long term interest rates diverged. This is pretty much the only point we tried to make here.

[10] The Clemens Critique: Ook hier geldt of je dit moet zien als een probleem of als de normale werking van een aanpassingsmechanisme (of alternatief of dit een goed criterium is om nominale convergentie te meten).
Our Reply: We agree that in a currency union, the only adjustment mechanism for asymmetric shocks is the fiscal budget. Nevertheless, the SGP intended to impose discipline and force member states to converge, at least in the long run. The data here show that this is very doubtful as there is no reversion to the mean and debt ratios remain high and diverge.

From the above we conclude that the nominal convergence criteria that were set in the Maastricht Treaty, whether intentional or not, have not secured the same level of nominal convergence since the introduction of the Euro. Inflation differentials remained persistent and did not cause real appreciation in the countries that grew fast or were more advanced. Long term nominal interest rates remained converged up to the Great Financial Crisis and ensuing Eurocrisis, arguably because risk premia failed to price in the no bail-out clause before the crisis and overreacted in the Eurocrisis, whereas fiscal discipline and consolidation also failed to bring Eurocountries closer together. We now turn to the more important real convergence.

Real convergence in the euro zone[11]

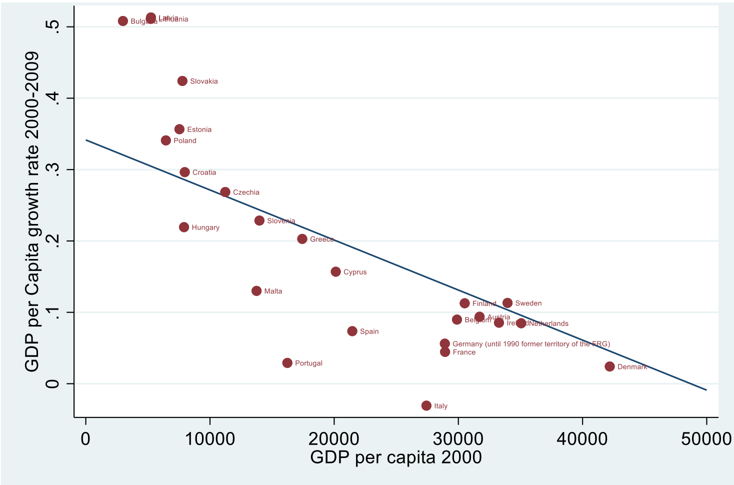
Looking at real convergence, different patterns between the South and the East emerge. In terms of beta-convergence (countries with lower GDP per capita levels catching up with the higher-level countries), convergence before the Maastricht treaty among the original (twelve) euro countries (North and South) was strong (IMF, 2018). This trend slowed down between 1993 and 2015, while there was only little beta-convergence between North and South in the first decade of the Euro. The differences between countries are significant: while Greece experienced strong growth, Portugal and especially Italy lagged. Figures 7 and 8 below plot the average growth rate against the initial GDP per capita for Eurozone countries for the first and second decades after the introduction of the Euro. The plotted trendline represents the linear regression line when all countries are included. Both decades show overall evidence of beta-convergence among the euro countries. However, when we zoom in on the Southern Member states, the picture is much different. The trendline would still be somewhat

---

[11] The Clemens Critique: Ik vraag met af of je het stuk over nominale convergentie echt nodig hebt in je verhaal. Uiteindelijk ben je geïnteresseerd in real convergence. Een korte alinea waarin je stelt dat nominale convergentie om verschillende redenen niet perfect kan zijn in een monetaire unie (aanpassing, risicopremies) en dat nominale en reële convergentie niet 1-1 gekoppeld zijn, zou een goede en voldoende aanloop kunnen zijn.
Our Reply: We agree with Clemens that we could have skipped the data and elaborate discussion of why nominal convergence has failed to materialize, but we are not convinced this would be sufficient for the critical reader.

downward sloping in Figure 7 but now slopes up in Figure 8.[12] This clearly shows that especially the Southern Member states not only failed to converge but started to diverge in and after the euro crisis.

Figure 7: Beta-Convergence 2000-2009



[12] The Clemens Critique: Ik vind het lastig om de boodschappen uit de verschillende grafieken te halen. Ook de eenheid op de verticale as is wat onduidelijk. Deze grafiek suggereert sterk dat er geen lineair maar een kwadratisch verband was in deze periode, met redelijk gelijke groei in de oude lidstaten en hogere groei in de nieuwe (Oosten).
Our Reply: We agree with Clemens that it takes some effort to see the different stories for South and East Europe. We do not agree that the relationship is quadratic. If one considers the Eastern Member states, they are arguably on a steeper line then the south, as we argue in the text, reflecting stronger convergence in 2000-2009 and still some convergence in 2009-2019. Only looking at the Southern Member states, we see the trendline go from negative to positive. These scatterplots do not show a causal link. Beta-convergence refers to the beta coefficient of the regression of average growth on initial GDP per capita, so a quadratic specification makes little sense.

Figure 8: Beta-Convergence 2009-2019[13]

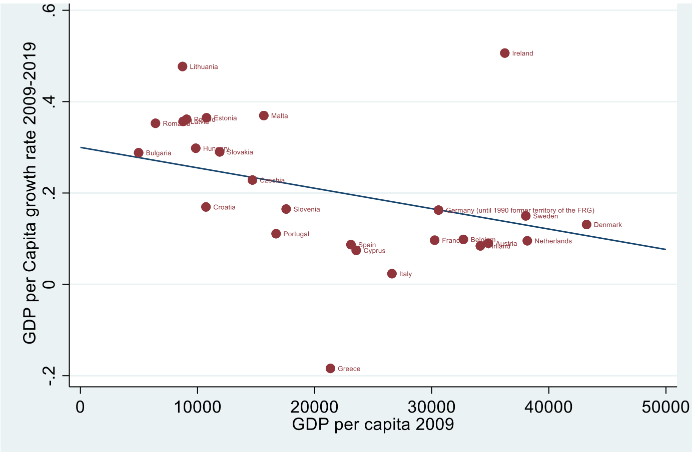
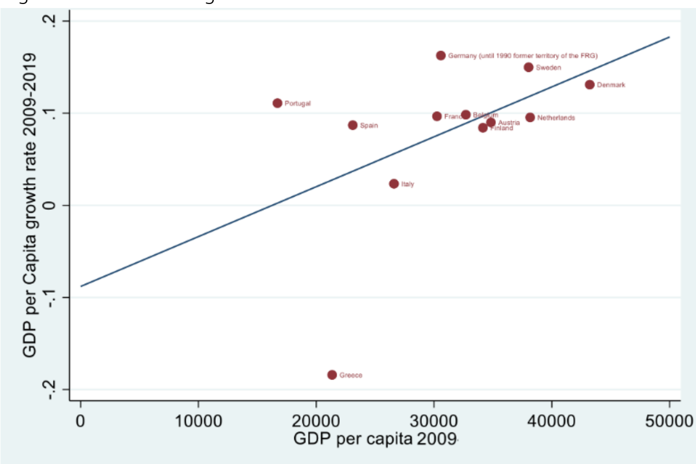


Figure 9: Beta-Convergence 2009-2019[14]

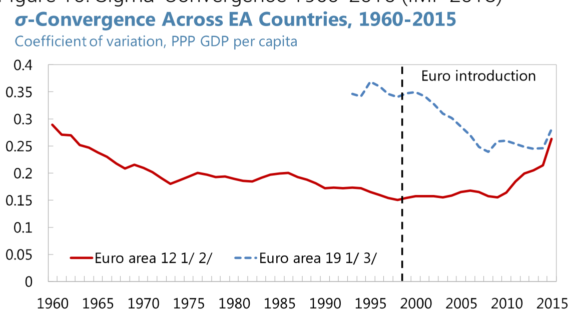


[13] The Clemens Critique: Ook hier kun je je afvragen of de lineaire lijn wel de goede boodschap geeft. Ierland en Griekenland zijn sowieso outliers.
Our Reply: Ireland and Greece are Eurozone countries. We do not think they are "outliers" because the measurements were incorrect. They are not typical countries, for sure, but you cannot drop observations because they do not fit a regression line. Taking them into account and estimating the sample average relationship between growth and initial level, we obtain the beta-coefficient. This is a linear line. The dots make it possible for the reader to critically assess our interpretation of the results, which are not much stronger than that the lines suggest a switch from con- to divergence in the South and continued but weaker convergence in the East. We believe that this conclusion survives eliminating Greece and Ireland.
[14] The Clemens Critique: Waarom hier 2008-19 ipv 2009-19 zoals in de vorige? Waar is Ierland nu? Als je Ierland en Griekenland eruit haalt, wordt de lijn denk ik vrijwel horizontaal.
Our reply: All figures are 2009-2019. We adjusted the text. Dropping Ireland and Greece would not make the line horizontal in Figure 8. Figure 9 gives the datapoints for only the Southern Member States and also without Greece the line would slope up.

However, compared to the rest of the EU, beta-convergence among the euro area members was significantly higher than between the EU countries in the period 1992 and 2015 (IMF, 2018).[15]

In the decades before the introduction of the euro, sigma-convergence (declining dispersion of GDP per capita levels) across the EU was visible, implying convergence of absolute income levels over time. For the original euro area countries, sigma-convergence after Maastricht slowed down and eventually stalled (IMF, 2018; see Figure 10).

Figure 10: Sigma-Convergence 1960-2015 (IMF 2018)



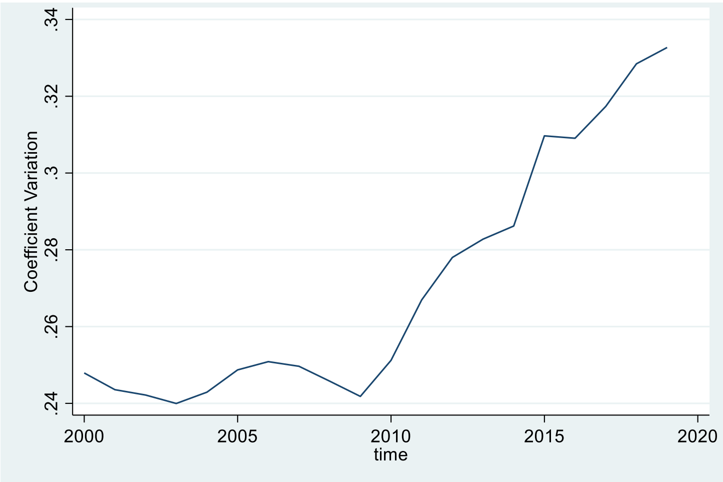
Sources: WEO database and IMF staff calculations.
1/ Excludes Luxembourg.
2/ Includes Ireland from 1970 and the Netherlands from 1980 onwards.
3/ Includes Lithuania from 1995 onwards.

The euro crisis reversed the convergence trend and caused diverging income levels among the Eurozone countries, notably between the North and the South. In Figure 11

[15] The Clemens Critique: Mijn take-away (uit jouw grafieken en andere literatuur) is dat de groei in de oude lidstaten nauwelijks samenhangt met initieel inkomensniveau (dus geen convergentie), en dat de groei in de nieuwe lidstaten hoger ligt (dus wel convergentie). Wat je hier niet ziet, is dat die hogere groeivoeten maar heel langzaam tot convergentie van inkomensniveaus leidt (hoge groei op laag inkomen loopt maar heel langzaam in op lage groei van hoog inkomen). Een tweede punt (ook voor de volgende sigma convergentie) is of je bij convergentie moet kijken naar GDP per capita of moet corrigeren voor verschillen in prijsniveaus (PPPadjusted GDP).
Our Reply: We agree that this should ideally be corrected for purchasing power parity but given what we know about prices (another reason to leave them in), we know that inflation differentials were small, persistent and positive, such that over time the purchasing power of the South would be eroded. In other words, the trends shown here are optimistic estimates of the real convergence. Clemens is of course correct that the core shows little convergence. As was shown in our toy-model, this is to be expected when countries are close to the steady state and shocks and productivity growth drive the dynamics between countries. We agree that the South and East converge, at least directly after introduction of the euro and the east continues to do so, while the South stops converging.

we show the trend in Sigma-convergence between North and South and for 2000-2019. The figure shows the clear upward trend in the coefficient of variation after 2009, when the euro crisis started to unfold.[16]

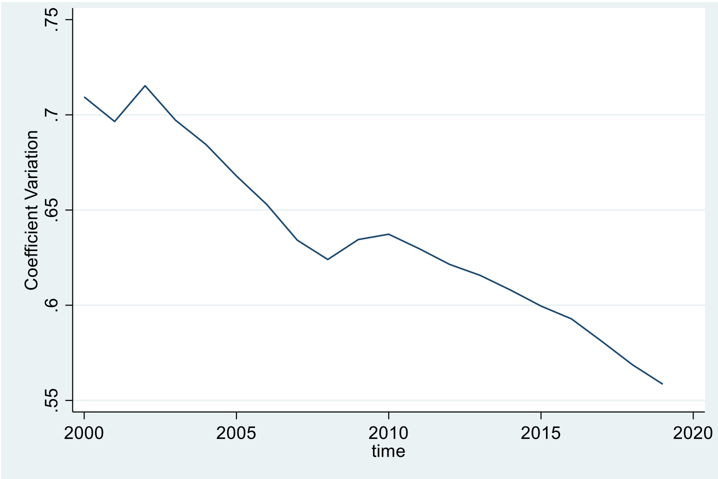
Figure 11: Sigma-Convergence North-South 2000-2019



In contracts, in Figure 12, we see the continued decline in the coefficient of variation between North and East, where the trend is consistently negative, with only a small blip when the euro crisis erupted.

[16] The Clemens Critique: Het zou wel nuttig zijn om een benchmark te hebben om het niveau te interpreteren. Is 0.25 hoog/laag ivm sigma convergentie binnen landen (VS, Duitsland, Spanje met een federale structuur)? In bovenstaande grafiek wordt weer weel PPP adj GDP gebruikt als ik het goed begrijp.
Our Reply: Indeed, we do not adjust for PPP and we focus in the figures on the trend, not the level, as it is hard to interpret what would be an appropriate benchmark for the Eurozone. Also, if we provided the levels for the USA, Germany or Spain, this would provide a very different benchmark. We therefore decided to focus on the trends only.

Figure 12: Sigma-Convergence North-East 2000-2019



Cyclical convergence in the euro zone

With respect to cyclical convergence, EMU countries have converged in the past two decades. While the business cycles of the future euro countries were already highly synchronized before the euro was introduced, this trend further increased after 1999 until the euro crisis (IMF, 2018, see Figure 13). During the euro crisis, the trend of cyclical convergence persisted.

Figure 13: Cyclical Convergence (IMF 2018)

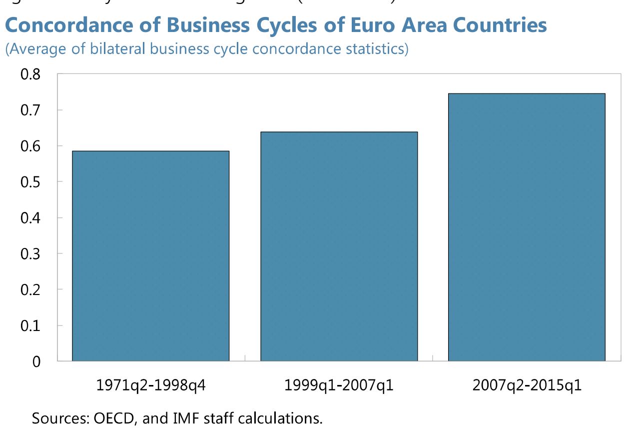
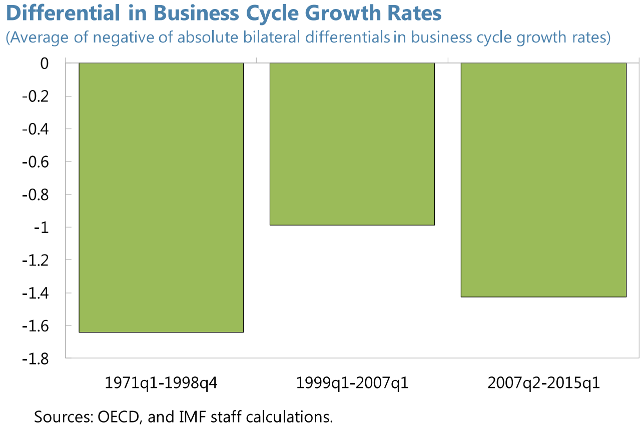
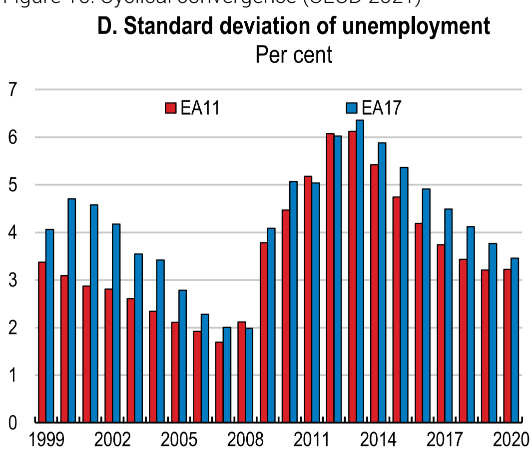


Figure 14: Cyclical Convergence (IMF 2018)



This is most likely because all euro members were hit by economic downturn and adopted a contractionary fiscal stance, deepening the recession. While business cycle convergence increased in the past three decades, the amplitudes of the business cycles in the euro countries show a different trend. Figure 14 displays that the reduction in differentials in business cycle growth rates of the first decade of the euro was reversed after the euro crisis hit.

Figure 15: Cyclical convergence (OECD 2021)



The consecutive business cycle convergence and divergence corresponds to unemployment patterns across the euro zone. Whilst unemployment levels converged in the first decade of the euro, they sharply diverged at the start of the euro crisis, impacting countries asymmetrically (Gori, 2021; see Figure 15). Since 2015, unemployment levels have started to converge again.[17]

Finally, financial cycle convergence shows a different pattern than business cycle convergence in the euro zone. The concordance of financial cycles decreased in the decade before the euro crisis. More precisely, Spain, Ireland, and Greece "experienced financial cycles of increasing duration and magnitude after euro introduction, in sharp contrast to core euro area countries, as cross-border bank flows from core country banks to the private (Spain, Ireland) and public (Greece) sectors boomed" (IMF, 2018). These differences in financial cycles (i.e., cheap credit flowing from the North to the South) was a major contributor to the euro crisis.[18] [19] We can conclude here that the euro has not hindered the cyclical convergence in the euro zone, but also has not prevented the backsliding that occurred when the euro was seriously tested. We conclude that the euro zone would benefit from cyclical convergence, as this reduces the need for adjustment and risk of asymmetric shocks. But to date, the EMU lacks the tools to ensure that gains in this respect are also robust.

What about productivity?

Productivity is an important driver of real convergence (ECB, 2017), and in fact, as we have seen in our toy model, the only driver that remains as countries approach their

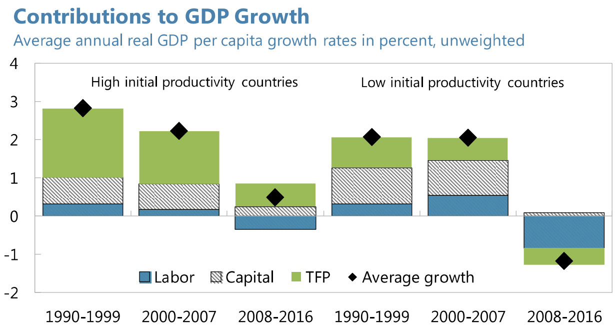
---

[17] The Clemens Critique: Er is behoorlijk wat extra literatuur over BC convergence/synchronization
Our Reply: Yes, we agree. But we believe this is sufficient for now. We are very open to your suggestions.
[18] The Clemens Critique: Wel erg kort door de bocht, Zeker een van de factoren.
Our Reply: We have adjusted the text to be less direct.
[19] This explanation of the euro crisis is broadly supported. See for example the "consensus narrative" https://voxeu.org/epubs/cepr-reports/rebooting-eurozone-step-1-agreeing-crisis-narrative and Stockhammer and Mohib Ali (2018), stating: "But since 1996 financial flows become more important as they were driven by property prices and stock prices. Overall, the picture that emerges is that real unit labour costs played a secondary role and the external imbalances in Eurozone arose as a consequence of strong domestic demand, spurred by credit boom in the South fuelled by the flow of private capital from the North and Anglo-Saxon countries."

steady state. In the past twenty years, productivity growth has been sluggish. Recently, productivity growth has been around 1% per year in the EU.[20]

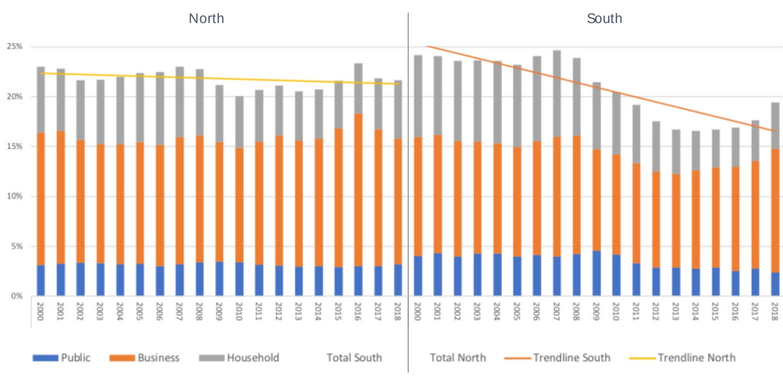
Significantly, the introduction of the euro did not lead to a productivity catch-up; the Total Factor Productivity (TFP) growth of countries with low initial productivity levels was lower than that of countries starting with high initial productivity and also stalled more sharply since the euro crisis (IMF, 2018; see Figure 16).

Figure 16: Sluggish Productivity Growth (IMF 2018)



Note: Productivity groups defined on the basis of labor productivity. Countries with high initial productivity include Austria, Belgium, Finland, France, Germany, Ireland, and the Netherlands. Countries with low initial productivity include Greece, Italy, Portugal and Spain. No 1990s data available for Austria.
Sources: AMECO, Haver Analytics, and IMF staff calculations.

Figure 17: Investment in Europe (% of GDP, SFL 2020)



North: Austria, Belgium, Germany, Ireland, Luxembourg, the Netherlands, Finland
South: Greece, Spain, France, Italy, Portugal

[20] https://publications.jrc.ec.europa.eu/repository/handle/JRC119785

The fall of productivity during the euro crisis of the Southern countries corresponded to rising unemployment and a sharp decrease in investments, mainly on the public and household level as visible in Figure 17.

One way of explaining the stalled productivity growth in these countries is the so-called "non-convergence trap", which states that "if an economy does not progress from growth driven by accumulation of capital to growth led by innovation, then it stops converging towards the technology frontier" (ECB, 2017).

Understanding divergence in the EMU[21]

Real convergence as discussed in the above builds on our toy model that predicts that regional disparities tend to decline in a process of market-driven convergence to a common technological frontier. There is also a different perspective on convergence, which argues that regional differences are likely to increase as the euro zone integrates its markets. This is explained by a positive feedback loop between a region's growth and the competitiveness of its export sector. Alexiadis (2019) explains this process as: "The mechanism by which the circular process of cumulative causation operates is captured in the relation between employment and output growth, known as Verdoorn's Law (Verdoorn, 1949) [...] Although workers move out of regions where employment opportunities are poor and move into regions where employment opportunities are good and wages are higher, there is now recognition that the consequences over the long-run are likely to be harmful to the origin regions. This is because migration tends to be age and skill selective with younger and more skilled or educated workers exhibiting greater mobility, thereby diminishing workforce quality in the origin region and making it less attractive to potential employers. Likewise, capital will flow into successful regions." Now that such arguments have also been raised to explain the rise of reactionary, populist right-wing politicians in

[21] The Clemens Critique: Volgens mij begint hier pas het punt dat je wil maken. Best wel heel laat in het paper en niet per se logisch vanuit de eerste 10 pagina's.
Our Reply: The paper is not making only one point. But we agree that here an important point follows from what went before.

especially the East, a European industrial policy, as proposed recently by Draghi, needs to carefully balance the need to boost productivity growth in Europe as a whole and ensure convergence among Europe's Member states to deliver on the promise of "Maastricht".

The fact that the decades following the introduction of a common currency did not bring real convergence across the Union, is, contrary to Northern preconceptions, not due to a lack of discipline or resolve in the Southern and Eastern Member States to abide by the rules or implement reforms. For example, Italy almost completely adopted the German labor market model in a series of reforms since the 1990s. Even more significant is the fact that Italy has displayed primary surpluses since 1992 with only two exceptions (2009 and 2010, at the start of the financial crisis; Beun, 2022). In a compelling article, Notermans and Piattoni (2021) contest the view that microeconomic reforms determine macroeconomic performance in a comparative study of Germany and Italy before and after euro implementation. They conclude that Italy suffered from a self-fulling prophecy dating back to their high debt legacy from the 1970s, which meant that a lack of trust of financial markets drove up interest rates and triggered the doom-loops that countries like Germany and the Netherlands could more easily avoid. At the same time, imbalances in the EMU, with productive Northern Member States achieving real depreciations with wage moderation and high domestic savings, hindered Italy to structurally strengthen its productivity and economy. The euro forced Italy to tackle current account imbalances with austerity and internal devaluations, depressing growth. Because of the practically unhindered export-oriented model of core countries such Germany and the Netherlands, Italy's productivity growth and thus real convergence remained vulnerable.

Fostering real convergence

As demonstrated in the above, the Maastricht promise of convergence has been only partially fulfilled. While the East experienced strong real convergence in the first decade of the euro, this trend slowed down in the second decade. Real convergence between North and South during the first ten years of the euro was weak, while there was divergence after the euro crisis started. The euro did not foster or force a productivity growth catch-up, and the Southern countries experienced decreased investments in the second decade of the euro. Nominal convergence criteria did not bring about real convergence, and it has been argued that the stringent fiscal rules and targets during the euro crisis led to untimely fiscal contraction. What remains is the conclusion that the euro has survived and by now has forced also the frontier countries to stop and reconsider. Sharing a currency implies sharing more than bills and a central bank. Asymmetric shocks continue to have a destabilizing effect that can affect all, whereas symmetric shocks have shown that a coordinated policy response is stronger than a single monetary policy alone. And while the reformed economic governance framework contains small improvements, a narrow focus on debt sustainability remains, limiting the fiscal room that countries need to make the investments to fulfil the goals of the Green Deal and new EU objectives.

The way forward towards a European investment agenda

In July 2020 the European Council agreed on a pan-European Covid Recovery Fund ("Next Generation EU") of 650bn linked to economic reforms. Centre piece of this temporary recovery instrument is the Recovery and Resilience Facility (RRF). Through this Facility, the European Commission raises funds by borrowing on the capital markets (issuing bonds on behalf of the EU). These funds are then made available in grants (359bn) and loans (291bn) to the EU member states in order to implement ambitious reforms and investments that a) make their economies and societies more sustainable, resilient and prepared for the green and digital transitions, and b) address the challenges identified in country-specific recommendations under the European Semester framework of economic and social policy coordination.

The corona crisis (2020-2021) was followed by a new crisis caused by Russia's invasion of Ukraine in February 2022. EU sanctions against Russia drastically limited the imports of cheap gas from Russia, thereby damaging European competitiveness. And the competitiveness of Europe had already been under pressure because China and the US have been making substantially higher investments in new innovative technologies during the past decade. Furthermore, after the inauguration of President Trump in January 2025 and his policy of imposing trade tariffs, European competitiveness received another blow, further underscoring the need for investments to foster the strategic autonomy of Europe.

In September 2024 the European Commission published the report "The Future of European Competitiveness". This report was prepared under the leadership of Mario Draghi, former President of the ECB. The report stipulates a need for a massive private and public investment, amounting to an additional 800bn *per year*, to enhance the competitiveness of the EU's economy. A substantial part of this amount should come from private investors, although some joint funding with the public sector of breakthrough innovation will be necessary.

Like the Covid Recovery Fund, it may be preferable to arrange the financing of investments in sustainable energy, digitalization and competitiveness, cross-border grids, and defense at the EU level. However, the Covid Recovery Fund implied financing at the EU level but investments made at the national level of EU member states. For several of the new challenges, this may also require careful consideration.

For strategic industries in Europe, the Draghi report advocates a coordinated EU strategy to bolster domestic production capacity and to protect key network structures. According to Draghi, there is a risk that a fragmented approach leads to weak coordination of priorities and demand requirements and lack of scale for domestic producers. To put this more bluntly, if we want to leverage the size of the EU domestic market and capitalize on economies of scale and scope, we must allow for agglomeration, clustering and, therefore, possibly divergence. A centralized EU budgetary allocation, provided it is well designed, could address these issues. In the case of pure public goods for the EU, such as a common defense and border security, a centralized allocation makes even more sense.

Such a centralized EU budgetary allocation would imply that not only the financing, but also the investment allocation decisions are taken at the EU level. An institution such as the European Investment Bank (EIB), which has been active since 1977, could then play an important role in the implementation of the Draghi investment agenda for Europe.

The preceding would mean that the procedure underlying the Covid Recovery Fund, financing at the EU level but investments made at the national level, should not be repeated. In a rather critical report, published in May 2025, the European Court of Auditors highlighted significant problems in terms of performance, accountability and transparency of spending under the Covid Recovery Fund. But more importantly, such an allocation through national member state budgets inevitably creates political pressure to distribute the funds "fairly" across all member states. The challenges for 2025-2050, however, dictate that we optimize our capital allocation for the Union as a whole.

An important aspect of the Draghi investment agenda is strengthening the productivity growth in the EU, including in member states that are lagging, thereby fostering real convergence. Improving the quality of institutions and governance in countries can also help to make productivity growth sustainable. But that goal is not achieved by treating very different regions equally.

In terms of EU policy, a European industrial strategy focused on productivity growth enhancing investments largely implemented by the EIB, and reforms compatible with the Green Deal's goals in weaker regions could lead the way. The latter would help to reshore European value chains in a more balanced equitable way across the EU. Fiscal policy, both in terms of rules and coordination, can help to create the right circumstances for successful industrial and investment policies. It is time to accept that Europe faces more common than national challenges and should start to behave as one to address them.

References

Alexiadis, S. (2019). Spatial-economic disparities and convergence. In *Handbook of Regional Growth and Development Theories* (pp. 436-450). London, UK: Edward Elgar Publishing.

Baldwin, R., Beck, T., Bénassy-Quéré, A., Blanchard, O., Corsetti, G., de Grauwe, P., Giavazzi, F., Gros, D., Kalemli-Ozcan, S., Micossi, S., Papaioannou, E., Pesenti, P., Pissarides, C., Tabellini, G., Weder, B., & Mauro, D. (2015). Rebooting the Eurozone: Step 1-agreeing a crisis narrative Introduction and summary. *VOX-EU Column*. https://cepr.org/voxeu/columns/rebooting-eurozone-step-1-agreeing-crisis-narrative

Barro, R. J., Sala-i-Martin, X. (2004). *Economic Growth*, Second Edition. Cambridge, MA: MIT Press.

Beun, H. (2022). *No more money to Italy*. *SFL-policy brief*. https://sustainablefinancelab.nl/nl/no-more-money-to-italy-reform-the-emu/. Utrecht, NL: Sustainable Finance Lab.

Creel, J. (2018). Convergence in EMU: What and How. In *Depth Analysis.* Requested by the ECON Committee. Brussel. Brussels: ECON Committee. https://www.europarl.europa.eu/RegData/etudes/IDAN/2018/614504/IPOL_IDA(2018)614504_EN.pdf

De Grauwe, P. (1996). Monetary union and convergence economics. *European Economic Review*, *40*(3-5), 1091-1101.

European Commission (2024). *The Future of European Competitiveness*. Report by Mario Draghi, September 9. Brussels: European Commission.

European Court of Auditors (2025). Performance-Orientation, Accountability and Transparency – Lessons to be Learned from the Weaknesses of the RRF, *Review* 02/2025, May 6. Brussels: European Court of Auditors.

Gori, F. (2021), Fostering cyclical convergence in the Euro Area, *OECD Economics Department Working Papers*, No. 1697, OECD Publishing, Paris, https://doi.org/10.1787/30cecd5a-en.

IMF (2011). A new action-based dataset of fiscal consolidation. *Working Papers* WP/11/128, Washington, DC: International Monetary Fund..

IMF (2018). Economic Convergence in the Euro Area: Coming Together or Drifting Apart?. *IMF Working Papers* WP/18/10, Washington, DC: International Monetary Fund.

Jones, C.I., & Vollrath, D. (2024). *Introduction to economic growth*. 4th ed., New York NY: W.W. Norton & Company.

Knot, K. (2020). *Samen sterker uit de crisis' Hoe we Europa weerbaarder, welvarender en duurzamer kunnen maken*. H-J-Schoo Lezing. https://www.dnb.nl/media/cmugjwfl/hj-schoo-lezing-klaas-knot-met-afbeeldingen-en-verwijzingen.pdf

Lucas Jr, R. E. (1976). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy 1*, 19-46.

European Central Bank, Ferdinand Heinz, F., Dorrucci, E., Diaz del Hoyo, J. L., & Muzikarova, S. (2017). *Real convergence in the euro area : a long-term perspective*, Frankfurt, DE: European Central Bank.. https://data.europa.eu/doi/10.2866/121513

Mankiw, N.G., Romer, D. & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 407-437.

Notermans, T. & Piattoni S. (2021). Italy and Germany: Incompatible Varieties of Europe or Dissimilar Twins? *German Politics 30* (3), 319-339. https://www.tandfonline.com/doi/full/10.1080/09644008.2021.1890717

OECD (2021), *OECD Economic Surveys: Euro Area 2021*, OECD Publishing, Paris, FR: Organisation of Economics Cooperation and Development (OECD). https://doi.org/10.1787/214e9f0a-en.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, *98*(5, Part 2), S71-S102.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, *70*(1), 65-94.

Stockhammer, E., & Ali, S. M. (2018). Varieties of Capitalism and post-Keynesian economics on Euro crisis. *Wirtschaft und Gesellschaft*, *44*(3), 349-370.

Van Tilburg, R. & Simić A. (2020). *The New Unconventional*. *SFL-policy brief.* Utrecht, NL: Sustainable Finance Lab. https://sustainablefinancelab.nl/nl/paper/the-new-unconventional/

Verdoorn, P J (1949). Fattori che regolano lo sviluppo della produttività del lavoro. *L'Industria 1*, 3-10.

Bayesian Economics, a Parable

Casper G. de Vries

Introduction

In this brief essay in honor of Clemens I like to touch on the use of Bayes' theorem in economics and econometrics. I was educated mostly as a macro-economist and frequentist econometrician, but later worked on non-conventional auctions using game theory. The article on biases in currency substitution estimates with Clemens, see Boon, Kool and De Vries (1988), has a Bayesian tinge if you like. Other than that, I did not actively use Bayesian econometrics. When I started thinking about this essay, I realized that I have nevertheless been a Bayesian in much of my research.

This is not because of the implicit use made of Bayes' theorem in the application to extreme value analysis. When estimating the tail index for financial data, most of the time the finance profession uses the Hill estimator, which is predicated on the assumption of heavy tails. This is more efficient, in the sense of lower variance, than using the DEdH estimator. The latter is applicable for all three types of limit laws (including exponentially thin tails and cases like the uniform distribution that have a finite upper bound). So implicitly one uses the common knowledge, build up from years of frequentist experience, that financial data are heavy tailed.

The other half of my academic work, however, has focussed on strategic decision making using game theoretic auction theory. This explains why I have been a Baysian all along. Much of the analysis of strategic decision making proceeds by using the concept of a Bayesian Nash equilibrium. In this essay I like to connect the two areas of game theory and econometrics around the theme of Bayes' theorem. In both areas one wonders how the prior distribution comes about, other than for computational convenience. Connecting the two areas provides some ideas on how economic insights bear on the type of priors one may want to use.

Bayesian Nash Equilibrium

In litigation two parties contest the ownership of a certain item, say (other applications are to tort law). Suppose that both parties value the item at v. If a dispute goes to court, asymmetric cases have already been settled out of court (assuming rational players who want to minimize the costs of hiring lawyers). So no party has a case that has intrinsically a higher merit than the opposing party. A judge has to decide to which party the item must be awarded based on the quality of the case brought by the lawyers. Here I make the despicable economist assumption that the quality of the case increases with the amount spent on ones' lawyers. The judge awards the case to the party that presents the best case. Given this simple production function, this means the party who has invested the most also wins. Then the question becomes how much to invest?

I first consider the case of American civil law in which contestants pay for their own legal expenses. Under continental law, the loser has to indemnify part of the winner's expenses, which is somewhat more involved due to this spillover effect. Let the expenses by the two parties be respectively x and y (which also identifies the two parties). So under American law party-x wins if $x > y$; in that cases it makes $v - x$. If, however, $x < y$, then party-x walks away with a loss of x due to the expenses on its lawyers.

It is easy to see there does not exist a pure strategy solution, as for any $k \in [0, v]$, the other party can win the case by investing just slightly more (at $k = v$, the opponent reverts to investing zero). One shows that this 'game' has just one solution which cannot be exploited by the opposing party and this is to pick an x randomly from the uniform distribution on $[0, v]$. Formally the strategic solution is equivalent to a so-called all-pay auction under complete information.

It is somewhat awkward to realize that since the two parties have equal merit, the outcome of the dispute is random. Moreover, one easily shows that in expectation the net gain from the court case is zero. For the interested reader, in the case that the losing party has to indemnify a percentage 1 – β of the winner's expenses under the continental rule, the Nash equilibrium strategy is again to expend randomly on one's lawyers. But the distribution is a bit more complicated:

$$F(x) = \frac{1}{1-\beta}\left\{1 - \sqrt{\frac{v}{v+2(1-\beta)x}}\right\} \qquad (1)$$

see Baye, Kovenock and De Vries (2012). In this case due to spillover effects, the litigating parties make less than zero in expectation:

$$-\frac{1-\beta}{2\beta}v \qquad (2)$$

A case of a negative sum game. This leads to the prediction that Americans more easily go to court, reckon the phrase: "see you in court". Europeans once entangled in a court battle fight harder to circumvent having to pay for the other parties' expenses. So the prediction is that less cases go to court in Europe, for which there is some empirical evidence.

In the above analysis it is assumed that each party is fully aware of the value v of the case. Now consider the situation of incomplete information. In this setting parties know the value of the case for themselves, but do not know how high the other party values the disputed item. Thus we speak of vx and vy as the valuations assigned to the case by party-x and party-y respectively. In the case of incomplete information the conventional game-theoretic approach is to assume that both parties have a common prior about the distribution of the valuation by the other party. But which prior to adopt? Auction and game theory is usually silent about this, though empirical analysis has to take a view. I now want to make the argument that the relationship between the case of complete and incomplete information can be of help.

Harsanyi (1973) has shown that there often exists a relationship between the complete information version of the setting in which there is a mixed strategy solution and the incomplete information setting with a pure strategy solution. For example, suppose that the production of evidence is quadratic in the expenses for lawyers (in the above the production function was linear).

With complete information the value of the case is v for both parties. The expected payoff to a contestant is

$$[v - x^2]\mathrm{Pr}\{\mathrm{win}\} + [-x^2]\mathrm{Pr}\{\mathrm{lose}\}$$
$$= [v - x^2]\mathrm{Pr}\{\mathrm{win}\} + [-x^2](1 - \mathrm{Pr}\{\mathrm{win}\}) \qquad (3)$$
$$= v\mathrm{Pr}\{\mathrm{win}\} - x^2$$

Assuming that the expected payoff is zero, yields:

$$\mathrm{Pr}\{\mathrm{win}\} = \frac{x^2}{v} \qquad (4)$$

Party-x wins with bid x if $\mathrm{Pr}\{\mathrm{win}\} = \mathrm{Pr}\{y \le x\} = F(x)$ and where $F(x) = x^2/v$ is the mixing distribution used by the opposing litigant. In a symmetric equilibrium, both parties independently draw their lawyer budgets from this simple beta distribution.
What is the pure strategy counterpart of this mixed equilibrium in the case of incomplete information? Suppose that the prior distribution of types that one may encounter in court is uniformly distributed on [0, w], i.e. with density 1/ w. In Baye, Kovenock and De Vries (2005) it is shown that the Bayesian Nash equilibrium under the American rule is to expend:

$$\mathrm{e}(\mathrm{v_x}) = \int_0^{\mathrm{v_x}} \mathrm{s}\frac{1}{\mathrm{w}}\mathrm{ds} = \frac{1}{2}\frac{\mathrm{v_x^2}}{\mathrm{w}} \qquad (5)$$

on the lawyers. A similar expression applies to the expenses by party-y.
Setting w = v /2 yields a quadratic distribution of $e(vx)$ on [0, v], given that litigant-x is unaware of which type the other litigant-y is. In this sense the pure strategy Bayesian Nash equilibrium expense (5) under incomplete information is related to the mixed strategy equilibrium (4) under complete information. Note that the prior distribution of types necessary for this relationship is specific and pinned down by the complete information version. Note also that in general the strategy in the incomplete information version with distribution of types $H(s)$ and associated density $h(s)$ is:

$$e(v_x) = \int_0^{v_x} sh(s)ds \quad (6)$$

The prior distribution can be any continuous distribution. But once the complete information version is specified, the prior is no longer a free parameter.

Thus in empirical work on litigation, one may first want to specify a density based on a frequentist analysis of the cost of producing legal evidence, mostly expenses on lawyer costs. Subsequently, one might then turn to the incomplete information case and specify a prior distribution of types based on the complete information analysis as is done above. Then one can test whether distribution of expenses fits. That is, in the above the predicted expense (6) generates a distribution that can in principle be rejected with a sufficient number of cases in the sample. So much for use of Bayes' theorem in strategic settings and possibilities for empirical analysis. Next is to discuss use of Bayesian econometrics.

Bayesian Econometrics

In econometrics I became convinced of the use of Bayesian econometrics when dealing with a stochastic process that may or may just not be a unit root process. Schotman and Van Dijk (1991) show the use of a diffuse prior in testing for a unit root as the distribution of the test statistic is smooth around the point where the process becomes a unit root process. This is different in the frequentist approach. The distribution of the test statistic on the first order autoregression parameter changes abruptly in going from a stationary AR(1) process to a unit root process. But in statistics one can almost always construct a local alternative which cannot be rejected given a certain amount of data. Hence, the Bayesain approach is natural (i.e., to let the data decide for which side there is more evidence).

In asset pricing Kozak, Nagel and Santosh (2020) recently made the argument that some priors are better than other priors in capturing the relevant characteristics that contribute most to the stochastic discount factor from a multitude of characteristics. Kozak et al. use principal components analysis and apply ridge regression. Their prior is such that it shrinks the contributions of low-variance components. The motive for selecting this prior is that the priors used in the literature produce unreasonably large Sharpe ratios for factors that contribute little to the risk premium. The posterior also gives more weight to high eigenvalue principal components and thereby helps to reduce overfitting. To achieve this, their estimator imposes a penalty on the model based maximum squared Sharpe ratio implied by the stochastic discount factor. The estimator is equivalent to minimizing the Hansen-Jagannathan (1997) distance and thus has a clear finance based motivation. So Bayesian econometrics can help to identify the more important drivers. Does this also apply to macroeconomics?

Bayesian econometrics has seen a huge increase in its use in macroeconomics since Smets and Wouters (2007) applied this approach to fitting a DSGE model for the Euro area. One may quibble about the size of the model, which is reminiscent of the of large scale macro models of the sixties that lost their charm due to their intractability. Nevertheless, Bayesian approach opened a new avenue for estimating complex models. The approach is nicely explained in the book by Herbst and Schorfheide (2016).

The first model considered by Herbst and Schorfheide (2016) is the textbook log-linearized three equation macro model from Gali (2008). The model has an IS-curve, a Phillips curve and a Taylor rule that captures monetary policy. In the presentation I use the notation by Herbst and Schorfheide for the coefficients, leave government expenditures aside, such that the specification is as in Gali (2008, ch. 3). The short run supply curve (the inflation rate π as a function of output gap y slopes upward with slope κ:

$$\pi_t = \beta E_t[\pi_{t+1}] + \kappa(y_t - \varepsilon_t) \qquad (7)$$

The long run aggregate supply curve runs vertical (which is nested in the short run curve by taking 1/ κ = 0); in which case only the supply shocks ε_t affect the output gap y_t = ε_t

If $1/\kappa = 0$ one obtains the classical dichotomy in which money is a veil (see e.g. Sargent, 1979).

The IS curve reads:

$$y_t = E_t[y_{t+1}] - \frac{1}{\tau}(i_t - E_t[\pi_{t+1}]) + \varphi_t \qquad (8)$$

Here i_t is the nominal rate of interest and so $i_t - E_t[\pi_{t+1}]$ captures the real rate of interest. A higher real rate deters investment and hence acts negatively on the output gap. Anticipated future positive movements in output already impinge positively on current spending and thus on demand through $E_t[y_{t+1}]$. Shocks are captured by stationary shocks φ_t.

The nominal rate of interest i_t is the instrument of the central bank. To close the model, suppose that the central bank follows a Taylor rule in setting the interest rate i_t according to:

$$i_t = \psi_1\pi_t + \psi_2 y_t + \upsilon_t \qquad (9)$$

Assuming no autocorrelation in the three shocks, the solution is well known as:

$$\begin{pmatrix}\pi_t \\ y_t\end{pmatrix} = \frac{1}{\left(1+\frac{1}{\tau}\psi_2\right)\frac{1}{\kappa}+\frac{1}{\tau}\psi_1}\begin{pmatrix}-\left(1+\frac{1}{\tau}\psi_2\right)\varepsilon_t + \varphi_t - \frac{1}{\tau}\upsilon_t \\ \varepsilon_t\frac{1}{\tau}\psi_1 + \frac{1}{\kappa}\left(\varphi_t - \frac{1}{\tau}\upsilon_t\right)\end{pmatrix} \qquad (10)$$

See e.g. Els and De Vries (2021). Thus inflation is:

$$\frac{-\left(1+\frac{1}{\tau}\psi_2\right)\varepsilon_t+\varphi_t-\frac{1}{\tau}\upsilon_t}{\left(1+\frac{1}{\tau}\psi_2\right)\frac{1}{\kappa}+\frac{1}{\tau}\psi_1} = \frac{-(\tau+\psi_2)\kappa\varepsilon_t+(\tau\varphi_t-\upsilon_t)\kappa}{(\tau+\psi_2)+\kappa\psi_1} \qquad (11)$$

We want to study the implications of the assumptions regarding the priors for the endogenous variables. Take for example the distribution of π_t implied by the prior distribution of the parameters $(\kappa, \tau, \psi_1, \psi_2)$, but conditional on the shocks $(\varepsilon_t, \phi_t, \upsilon_t)$. Focus on the part:

$$\frac{-\upsilon_t\kappa}{\tau+\psi_2+\kappa\,\psi_1} \quad (12)$$

Herbst and Schorfheide (2016) assume that the three parameters (τ, ψ_1, ψ_2) have independently gamma distributed priors, while κ has a uniform prior on [0,1]. For simplicity, set κ = 1 (the argument below carries over for the case where κ is uniformly distributed). To see the implication of these assumptions, suppose that all three (τ, ψ_1, ψ_2) have gamma densities:

$$f_x(x; z) = \frac{1}{\Gamma(z)} x^{z-1} \exp(-x) \quad (13)$$

with z > 1. Then it is well known that the convolution $x = \tau + \psi_1 + \psi_2$ has a gamma density

$$f_x(x; s) = \frac{1}{\Gamma(s)} x^{s-1} \exp(-x) \quad (14)$$

and where $s = z_\tau + z_{\psi_1} + z_{\psi_2}$ > 3 is the sum of the individual gamma density parameters (as the z's exceed 1). Consider (12), we need that distribution of y= $1/x$. In the end we only need the weaker assumption s > 1.

Claim 1:
Suppose s > 1. Then the distribution of
$y = 1/(\tau + \psi_1 + \psi_2)$

has a power upper tail.

Proof. If x has a gamma density

$$f_x(x; s) = \frac{1}{\Gamma(s)} x^{s-1}\, exp(-x)$$

Then y = 1/ x has density (by a transformation of variable)

$$f_y(y; s) = \frac{1}{\Gamma(s)} y^{-s-1}\, exp(-1/y)$$

Consider the regular variation properties of the density of y:

$$\lim_{t\to\infty}\frac{f_y(yt;s)}{f_y(t;s)}=\lim_{t\to\infty}\frac{(ty)^{-s-1}exp\left(-(ty)^{-1}\right)}{t^{-s-1}exp(-1/t)}=(y)^{-s-1}\lim_{t\to\infty}\frac{exp\left(-(ty)^{-1}\right)}{exp(-1/t)}=y^{-s-1}$$

which is a regularly varying function with tail index s + 1 > 0 as s > 1 by assumption.

Because with s > 1, the density of x has zero in its support, even if the probability of realizing zero has probability zero. Then outcomes close to zero yield extremely large values of $1/y$. This makes that the density of y follows a power law in the tail. If the three distributions of (τ, ψ_1, ψ_2) all have z > 1, then the distribution of y has at least a finite mean, variance and skew. But the higher moments will be unbounded.
Suppose that the shock to the Taylor rule υ is normally distributed, as is assumed in Herbst and Schorfheide (2016). If the distribution of y has a power upper tail, then one can show that υy also has a distribution with a power upper tail (use Breiman's lemma). The implication of this is that the endogenous variable π has a distribution with heavy tails. Whether the priors (for X and Y) were selected intentionally for this reason was not directly clear to me, but inflation data do have a tendency to be heavy tailed.

I have a few other observations on the priors chosen by Herbst and Schorfheide (2016) and their internal consistency. For the Gali model it is assumed that the prior for the slope of the Phillips curve κ is uniformly distributed. While τ, which captures the slope of the Phillips curve 1/ τ, is gamma distributed. However, in their derivation $\kappa = c.\tau$ and where c is a constant. So the assumptions regarding the priors are not mutually consistent, a form of model inconsistency.

The shock to the monetary policy rule should also be very different from the specification of the residual in the IS curve. In the latter case the central limit theorem pointing to a normal distribution appears reasonable, given the many shocks that impinge on the aggregate demand curve. But for shocks to the Taylor rule, this may be different as these come from a monopolist. So economic arguments can be used to improve the selection of priors.

Closing Remarks

In this essay I have tried to use economic theory to whittle down on the sheer unlimited choice of priors. Priors should come from economic insights. This implies, as explained in both parts of the essay, that one is not 'free to choose'. The priors should be model consistent. This is like in rational expectations that imposes consistency of expectation formation with the underlying economic model. The least requirement is that priors are coherent with other priors. Van Dijk (2024) suggests other avenues for future work in Bayesian econometrics. Regarding future developments in game theory and Bayes' theorem, perhaps coupled with strategic games played by central banks, might be a novel area to explore for Clemens, once he has retired. If he ever does, but I leave the priors for this state of affairs to the reader.

References

Baye, M.R., D. Kovenock and C.G. de Vries (2005). 'Comparative analysis of litigation systems: An auction-theoretic approach', *The Economic Journal, 115*, 583-601

Baye, M.R., D. Kovenock and C.G. de Vries (2012). 'Contests with rankorder spillovers', *Economic Theory, 51*, 315-350

Boon, M., C. Kool, C. de Vries (1988). 'Simulating currency substitution bias', *Economic Letters 28*, 269-272

Els, P. van and C.G. de Vries (2021). The Phillips curve, stable or rotating slopes, in J.M Berk and K. Knot, *Chief Economist, essays in honor of Job Swank*, DNB, ch.4

Gali, J. (2008). *Monetary Policy, Inflation and the Business Cycle*, Princeton University Press, New Jersey

Hansen, L.P. and R. Jagannathan (1991). 'Implications of security market data for models of dynamic economies', *Journal of Political Economy 99*, 225–262

Harsanyi, J. (1973). 'Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points', *International Journal of Game Theory 2*, 1-23

Herbst,E.P. and F. Schorfheide (2016). *Bayesian Estimation of DSGE Models*, Princeton University Press, New Jersey

Kozak, S., S, Nagel and S. Santosh (2020). 'Shrinking the cross-section', *Journal of Financial Economics 135*, 271-292

Sargent, T.J. (1979). *Macroeconomic Theory*, Academic Press, New York

Schotman, P. C., and H. K. Van Dijk (1991). 'A Bayesian analysis of the unit root in real exchange rates', *Journal of Econometrics, 49*, 195-238

Smets, F. and R. Wouters (2007). 'Shocks and frictions in US business cycles: A Bayesian DSGE approach', *American Economic Review, 97*, 586608

Van Dijk, H.K. (2024). 'Challenges and Opportunities for Twenty First Century Bayesian Econometricians: A Personal View', *Studies in Nonlinear Dynamics & Econometrics 28*, 1-22

Competition among Dutch pension funds: is there any?

Jacob A. Bikker, Michiel van Leuvensteijn and Jeroen J. Meringa

Introduction

"Men who are anxious to win the favour of a prince nearly always follow the custom of presenting themselves to him with the possessions they value most, or with things they know especially please him." (Machiavelli, The Prince) To you, Clemens, we provide our (limited) knowledge of measuring competition. Of course, we leave it to you whether this must be considered as our most valuable asset or whether we are under the illusion that this topic will please you most.

Competition in markets is ultimately about the welfare that markets create for humanity. Markets allocate resources to firms that produce goods and services. The more efficient this allocation is, the more welfare will be created. This allocation process can be influenced by structural elements and the conduct of firms. Important structural elements are entry and exit barriers, the number of firms in the market, and transaction and information costs. On the conduct side, firms may try to eliminate their competitors by, for example, taking over upstream firms that provide their competitors with essential goods.

Measurement of competition is a complex matter as competition is not directly observable. The literature provides different strands that vary in complexity and theoretical underpinning (Bikker and Bos, 2008, Bikker and Van Leuvensteijn, 2014, Bikker and Spierdijk, 2017). One strand is focused on structural measures of competition and another on non-structural measures. Examples of structural measures are static concentration measures, and dynamic measures, such as entry and exit rates, and the Structure-Cost-Performance (SCP) model. Non-structural measures of competition are based on the conduct of firms. Examples are indicators developed by the New Empirical Industrial Organisation literature (NEIO). The most applied performance measures are price mark-ups (Lerner index, Elzinga and Mills, 2011; Giocoli, 2012), correlations of output prices with input costs (H-statistic, Panzar and Rosse, 1982, 1987) and the conjectural variation parameter λ of the Bresnahan (1982) or Lau (1982) models.

This chapter draws on the framework of Hay and Liu (1997) and Boone (2008), who show that when substitutionability of products increases, or when entry barriers diminish, the elasticity between profits and efficiency increases. In a more competitive market, therefore, inefficient firms are taken over by more efficient firms, are pushed out of the market or have to adjust and become more efficient so that the relationship between efficiency and profits is strengthened. This measure of Hay and Liu and Boone (elasticity between profits and efficiency) has the advantage of having a clear theoretical underpinning which is directly related to the efficient allocation of resources in a market. Further, it uses little data compared to alternatives such as the Bresnahan model.

The main contribution of this chapter is its application of this measure to pension funds. Bikker and Van Leuvensteijn (2008) investigate this measure for life insurance companies, and Van Leuvensteijn et al. (2011) apply it to the loan markets of banks. Life insurance companies and banks operate in markets in which risk is highly regulated: life insurance companies are subject to the Solvency II framework and banks to the Basel III Accord. Risk regulation is less intensive for pension funds, where – within limits – the risk profile varies according to their risk appetite, depending, among others, on the average age of their members.

More than for life insurance products, the product of pension funds is very homogenous, i.e. lifelong pension benefits for their members. In principle, this makes it easier to measure competition. However, competition between pension funds in the Netherlands is severely limited because employees of companies (in the case of the company pension funds) or industries (in case of industrywide pension funds) cannot choose between pension funds, the so-called great obligation ('grote verplichtstelling'). Still, competition between pension funds may occur when company pension funds merge or are taken over, or when industrywide pension funds are not mandatory for firms (when they are mandatory, we speak of small obligation, 'kleine verplichtstelling'). Furthermore, pension funds compete for the best asset managers. Up to a certain level, our approach also may reflect competition between pension funds service providers, such as pension administration offices and external investment companies, as their costs are included in the costs that pension funds report ("look through principle").

Our analysis of competition among pension funds is to our knowledge the first of its kind, and the level of competition has thus far been unknown. Our ex ante expectation was, of course, that the pension industry competition level is lower than in the life insurance industry, because there is no competition for customers. Given this lack of competition for customers the result could well be that competition is absent altogether and the mechanisms for competition described above are too weak to measure. In this sense, this chapter is an experiment.

The structure of this chapter is as follows. Section 2 provides a literature overview on competition measurement, section 3 sketches the pension system in the Netherlands, the following three sections present the model, the data used, and the empirical results. Section 7 presents our conclusions.

Literature overview.

The Structure Conduct Performance (SCP) framework is the basis of most of the measures of competition on a market level. The idea is that structural properties of the markets define the conduct of firms in terms of price and quantity decisions and result in the performance of the markets in terms of profits, profit margins or market share. Structural measures of competition include the level of concentration in the market, entry and exit barriers and transaction and information costs, for instance. The Herfindahl-Herschman index of concentration is often used in the competition literature. The New Empirical Industrial Organisation literature also focuses on the conduct of firms considering non-structural measures like the Lerner index (Genesove and Mullin, 1998) and the divergence between price and marginal revenue, the Bresnahan measure (Bresnahan, 1982, and Lau, 1982). Another example of measuring conduct is the H-Statistic as developed by Panzar-Rosse (1987). In this measure, the transmission of costs to prices is central. Markets in which changes in costs are strongly related to changes in prices are more competitive than markets in which the pass-through is sluggish. In competitive markets, firms are price takers with little pricing power.

A last strand of literature is devoted to measuring the X-efficiency of firms. X-efficiency reflects the managerial ability to reduce costs, controlled for production volumes and input prices. X-efficiency is defined as the difference in costs between a firm and the best practice firms of similar size and input prices (Leibenstein, 1966). The expectation is that competition in the market would diminish the X-efficiency of firms, i.e. the degree of efficiency maintained by firms under conditions of imperfect competition.

This chapter uses the indicator based on the framework of Hay and Liu, and Boone which focuses on the relationship between profits and efficiency. The idea is that in a more competitive market, the relationship between profits and efficiency is stronger because inefficient firms are pushed out of the market, are taken over by more efficient firms or adjust and become more efficient. In all these cases, the relationship between profits and efficiency becomes stronger. Boone (2008) shows that there is a monotonic positive relationship between relative efficiency and relative profits. In Bikker and Van Leuvensteijn (2014) we renamed this approach the Performance Conduct Structure (PCS) indicator: efficiency (or profit) in a competitive market leads to a change in its structure as more efficient firms grow. It is the opposite view of the well-known Structure Conduct Performance (SCP) framework, which postulates causal relationships between the structure of a market, the conduct of firms in that market and their economic performance.

The advantage of this PCS indicator is that it requires relatively little data compared to, for example, the Bresnahan model. Furthermore, this measure does not suffer from the shortcomings of the Lerner index. In general, more competition may reduce the Lerner index as it reduces the price-cost margin, and a low Lerner index would indicate high competition. Competition may result in more efficient firms where some of them may decide to have a higher price cost margin, which is possible due to their lead in efficiency. Finally, Corts (1999) criticized the (elasticity adjusted) Lerner index because efficient collusion could not be distinguished from Cournot competition.

The PCS indicator has been used for various segments of the financial sector. Bikker and Van Leuvensteijn (2008) implemented it for the life-insurance market, Van Leuvensteijn et al. (2011) applied this measure to the loan markets of Europe, US and Japan, while Bikker (2017) and Bikker and Bekooij (2024) adopted it to the health insurance industry. For the pension fund industry, measuring the X-efficiency provides an initial indication of market competitiveness. Alserda et al. (2018) show large economies of scale for pension fund administrations, but modest economies of scale for investment activities, and find that many pension funds have substantial X-inefficiencies for both administrative and investment activities.

The pension system in the Netherlands

The institutional structure of the Dutch pension system is made up of three pillars, similar to the situation in most other developed countries. The first pillar consists of a public pension scheme and is financed on a pay-as-you-go basis. It offers a basic flat-rate pension to all retirees and aims to link the benefit level to the legal minimum wage. The pension benefit age moved gradually from 65 years until 2012 to 67 years in 2024 (dynamically linked to life expectancy). The second pillar provides former employees with additional income from a collective, contribution-based supplementary scheme. The prescribed pension age is 68 years. The third pillar is composed of tax-deferred personal savings, which individuals undertake at their own initiative and expense. The supplementary or occupational pension system in the Netherlands is typically organised as a funded defined-benefit (DB) or collective defined contributions (CDC) scheme. The benefit entitlement is determined by years of service and a reference wage, which in more recent years has been linked to wages over the years of service. The second pillar takes the public scheme benefits into account, while the third pillar's tax deduction takes the sum of the benefits from the first two pillars into account. Supplementary schemes are usually managed collectively by pension funds. Three types of pension funds exist. The first is the industry pension fund, which is organised for a specific industry sector (e.g. construction, healthcare, transport). Participation in an industry pension fund is mandatory for all employers in the relevant sector, with a few exceptions. An employer may opt out if it establishes a corporate pension fund that offers a better pension plan to its employees. Where a supplementary scheme is agreed by employers and employees, managed by either a corporate pension fund or an industry pension fund, employee participation is mandatory, governed by collective labour agreements. The third type of pension fund is the professional group pension fund, organised for a specific group of professionals, such as the medical profession or notarial profession. The Dutch pension fund system is comprehensive. In 2022, almost 89% of the employees are covered, but self-employed people need to arrange their own retirement savings. In that year, total pension fund assets in the Netherlands amounted to some €1,512 billion, or 158% of GDP, ranking the Dutch pension system, in terms of the assets-to-GDP ratio, as the largest in the industrial world. The government, employees and employers have agreed to transform the pension system into a kind of defined contribution system, resulting into the Future Pensions Act. This law revises the Dutch pension system and came into effect on 1 July 2023. Pension funds currently have until 2028 to switch to the new system. This system may have a collective buffer to soften setbacks, at the choice of the pension funds.

The Performance Conduct Structure (PCS) model

The theoretical PCS model

The PCS model is based on two notions. First, more efficient firms, i.e., firms with lower marginal costs, gain higher market shares or profits. Second, this effect is stronger in more competitive markets (Van Leuvensteijn et al., 2011). The PCS indicator is the empirical operationalisation of a theoretical model developed by Boone (2008) who shows that there is a continuous and monotonically increasing relationship between profit differences and the level of competition if firms are ranked by decreasing efficiency. In other words, there is a negative relationship between efficiency, measured in terms of marginal costs, and profits. The more intense this negative relationship is, the more competitive markets will be.

So, in practice, a negative sign between marginal costs and profits is expected. This elasticity is called the PCS indicator The fact that this relationship is both continuous and monotonic is the main advantage over more traditional measures of competition such as the H-statistics and Lerner index. Another advantage is that the PCS indicator is not dependent on assumptions about the type of competitive model, as in the cases of Bertrand or Cournot.

Following Boone (2008) and Xu et al. (2016), we consider an industry where each firm *i* produces one product q_i (or portfolio of products), which faces a demand curve of the form:

$$p(q_i, q_{j\neq i}) = a - bq_i - d\sum_{j\neq i} q_j \qquad (1)$$

and has constant marginal costs mc_i. This firm maximises profits $\pi_i = (p_i - mc_i)q_i$ by choosing the optimal output level q_i. We assume that $a > mc_i$ and $0 < d \leq b$.[1] The first-order equilibrium condition for a Cournot-Nash equilibrium can then be written as:

$$a - 2bq_i - d\sum_{j\neq i} q_j - mc_i = 0 \qquad (2)$$

[1] d ≤ b follows from the assumptions of Boone (2008), where he shows how a decline in entry costs strengthens the relationship between efficiency and profits.

When *N* firm produce positive output levels, we can solve the *N* first-order conditions (2), where mc_i = $1/\eta_i$, yielding:

$$q_i(\eta_i) = \frac{(2b/d-1)a-(2b/d+N-1)(1/\eta_i)+\sum_j(1/\eta_j)}{(2b+d(N-1))(2b/d-1)} \quad (3)$$

The argument of function q_i, η_i, is efficiency, the inverse of marginal costs. We define profits π_i as variable profits excluding entry costs ε. Hence, a firm enters the industry if, and only if, $\pi_i \geq \varepsilon$ in equilibrium.[2] Note that Eq. (3) provides a relationship between output and marginal costs. From $\pi_i(mc_i) = (p_i - mc_i)q_i$, it follows that profits depend on efficiency in a quadratic way, i.e.

$$\pi_i(\eta_i) = \frac{(2b/d-1)a-(2b/d+N-1)(\frac{1}{\eta_i})+\sum_j(\frac{1}{\eta_j})}{[(2b+d(N-1))(2b/d-1)]}\left(p_i - (\frac{1}{\eta_i})\right) \quad (4)$$

The Relative Profit Differences is then defined as $RPD = \frac{\pi(\eta^{**})-\pi(\eta)}{\pi(\eta^{*})-\pi(\eta)}$ for any three firms with $\eta^{**} > \eta^{*} > \eta$. The Relative Profit Differences provides the relative profits of a firm with efficiency η^{*} to the profits of the most efficient firm with efficiency η^{**} and the profits of the least efficient firm with efficiency η. In this market, competition can increase in two ways. First, competition increases when the services produced by the various firms become closer substitutes, that is, *d* increases (keeping *d* below *b*). Second, competition increases when entry costs *ε* decline. Boone (2008) proves that RPD is an increasing function of interaction among existing firms ($\frac{dRPD}{dd} > 0$) and a decreasing function of entry costs ($\frac{dRPD}{d\varepsilon} < 0$). In other words, RPD increases when competition intensifies, i.e., fiercer competition increases profits of more efficient firms by larger amounts than those of less efficient firms.

[2] The model does not have a time dimension, so as long as the profits in the considered time period is higher than the entry costs, a firm will enter.

The empirical PCS model

Profits

Following Xu et al. (2016) and Bikker and Van Leuvensteijn (2008), we adjust for the particular nature of the profits of pension funds. The relationship between profits and marginal costs in logarithms is central. However, the model must be adjusted in two ways. First, profits from asset management are strongly affected by risk. The more investment risk is taken, the more return on assets can be expected, on average. This relationship is not considered in the original PCS indicator model. Therefore, the volatility of the investment portfolio of the pension funds is included in the PCS equation. Second, we define the marginal costs as the investment management costs expressed as a percentage of assets.[3] We assume that management costs should be as low as possible, the traditional interpretation of cost as 'waste'. Of course, additional costs may lead to better investment decisions, i.e. the return on management costs may be non-linear. We treat performance fees as a separate class of costs. These costs are costs that materialise after profits are known and should be treated separately from other costs that have to be taken in advance. They indicate co-investments by private equity parties and are therefore strongly related to the risk profile of pension funds. We expect a different – possibly positive – relation with profits. The model becomes:

$$\ln(\text{profits}_{it}) = c + \beta \ln(\text{mc}_{it}) + \alpha \ln(\text{perff}_{it}) + \zeta \ln(\text{risk}_{it}) + u_{it} \qquad (5)$$

where profits is the net return on assets in euro, mc is marginal investment management costs, perff are performance fees, and risk is the volatility of the portfolio, for i=1,…,N firms and t=1,…,T periods. The key hypothesis to be tested is: $\beta = 0$ against the alternative $\beta < 0$. We split the sample period into two subperiods. An increase over time of β, in absolute terms, would indicate an increase in competition.

3 For practical reasons, we ignore constant costs. In Bikker and Van Leuvensteijn (2008) we show that the effect of ignoring on the estimations is negligible.

As a robustness test, we also estimate a model with cross-terms to take correlations between the explanatory variables into account. In this model the first derivative varies between pension funds, but the competitiveness of the average pension fund can be derived. The model is as follows:

$$\ln(\text{profits}_{it}) = \beta \ln(\text{mc}_{it}) + \alpha \ln(\text{perff}_{it}) + \zeta \ln(\text{risk}_{it}) + \gamma_1 \ln(\text{perff}_{it}) \ln(\text{risk}_{it}) + \gamma_2 \ln(\text{risk}_{it}) \ln(\text{mc}_{it}) + \gamma_3 \ln(\text{perff}_{it}) \ln(\text{mc}_{it}) + c_t \qquad (6)$$

Market shares

Most applications of the PCS model are performed on market shares instead of on profits (e.g. Hay and Liu, 1997). Apart from a constant, profit is the product of market shares and profit margins. More efficient firms may increase their market shares through lowering their prices, that is, by passing on a part – or the whole – of their efficiency lead to their (potential) customers. Increasing market shares of efficient firms is a clear signal that competition functions well. The role of profit margins is more complicated. Fierce competition may force efficient firms to pass on all their cost advantage to their (potential) customers, at the cost of their profits.

The PCS model for market shares reads as:

$$\ln(\text{market shares}_{it}) = \beta \ln(\text{mc}_{it}) + \alpha \ln(\text{perff}_{it}) + \zeta \ln(\text{risk}_{it}) + c_t \qquad (7)$$

The variable marginal costs is approximated by either the sum of investment and administrative cost margins or by each of the two components of this cost margin.

Data used

This chapter is based on unique and extended quarterly reports of 280 pension funds over 2012Q1–2023Q1 to their prudential supervisor, De Nederlandsche Bank (DNB). Pension funds with less than one hundred members (in 2022 only four funds) and funds which did not report all the items needed for our research have been excluded. At year-end 2022, the selected funds together managed 97.9% of the total pension investments of €1,433 billion. Key variables are net returns, investment costs and risk.

Table 1 presents the data for the entire period, as well as two sub periods, in order to show developments over time.

Average net returns over pension funds and time are expressed in annual figures. Over the entire sample period that return is 5.9%, higher in the earlier years (8.7%) compared to the later ones (2.9%). In our central log-linear model we use only positive net returns, a subsample of around 70%, but we release this restriction in one of the robustness tests. Investment costs are, on average, 0.56%, falling over time from 0.61% to 0.50%. Investment costs are split into three components: management costs, performance fees and transactions costs. Management costs fall over time (from 0.46% to 0.39%), where performance fees increase (from 0.10% to 0.14%). Annual administrative costs also decline, from 0.41% to 0.29%. As the number of pension funds decrease rapidly (from 270 to 206), particularly due to the fall in the number of company funds, the average market share increases (from 0.37% to 0.49%). Volatility is defined as the average percentual change in quarterly returns. For the entire investment portfolio, it was 6.0% over the full sample period, hardly changing across the two subperiods. Finally, we present the asset allocation across investment categories. We observe a slight increase of risky assets over time.

Table 1. Survey over our data set

	2012Q1-2023Q1	2012Q1-2017Q2	2017Q3-2023Q1
Avg annual net returns over pension funds and time (mln euro)	293	355	221
Idem, in %	5.94	8.71	2.59
Avg annual positive net returns over pension funds and time (mln euro)	770	509	1,033
Avg annual negative net returns over pension funds and time (mln euro)	-998	-505	-1,380
Avg annual management costs over funds and time in %	0.43	0.46	0.39
Idem, performance fees in %	0.10	0.10	0.10
Avg investments costs over pension funds and time in %	0.56	0.61	0.50
Avg market shares based on total assets over pension funds and time in %	0.42	0.37	0.49
Avg annual administrative costs over funds and time in %	0.34	0.41	0.27
Avg number of pension funds over time in %	237	270	206
Avg share of industry sector pension funds in %	20	17	22
Avg share of company pension funds in %	71	75	67
Avg share of professional group pension funds in %	4	3	4
Avg quarterly volatility over pension funds and time in %	5.95	5.88	6.04
Avg volatility fixed income in %	7.76	9.27	6.04
Idem shares	10.81	8.38	12.59
Idem real estate	8.32	5.48	10.20
Idem private equity	8.52	7.28	9.45
Idem hedge funds	9.76	8.51	11.89
Idem commodities	14.49	12.26	18.28
Avg share of fixed income over pension funds and time in %	62.02	63.10	60.73
Idem shares	30.34	29.43	31.40
Idem real estate	6.21	5.94	6.51
Idem private equity	2.78	2.65	2.88
Idem hedge funds	2.90	2.98	2.71
Idem commodities	2.88	2.95	2.74

Source: DNB, own calculations. Since 2015, private equity has been called an alternative investment, due to a minor change in the definition.

Empirical results

Profits

We apply our PCS competition model from Eq. (5) to pension fund net profits, that is positive net returns, from our data set over 2012Q1-2023Q1. We observe positive returns in more than 70% of the quarters. Roughly speaking, returns are positive in quarters where the share index increases. Of course, this data selection may affect the results, therefore, we also apply two alternative approaches which include all observations. We present clustered standard errors (CSE), where each pension fund defines a cluster. Estimation results are presented in Table 2, upper panel. All explanatory variables have highly significant coefficients with signs as expected. Management cost, the waste type cost, has a negative coefficient at -0.48 (first panel, full period), which is in line with the PCS model theory: the lower pension funds' costs, the higher their net returns. This is not the direct costs effect, as the returns are taken net of costs. Hence, this negative cost effect describes typically an increase in market shares, as we will also demonstrate below in the next Section "Market Shares". Efficiency contributes to pension fund grows because of some pressure, possibly of competition, despite the limited competitive pressure in the Dutch pension market, or principal agent types of relationships, such as checks and balances in the pension market and pressure of participants instead of shareholders. Poorly performing pension funds may have higher probability of being taken over by better performing pension funds. Industry sector funds which underperform may lose their mandatory status, enabling companies in that sector to transfer their pension capital to another pension fund. We observe that the strength of the costs-returns relationship, at - 0.48, is weak, as this value is smaller in absolute terms than in other financial industries such as banking and insurance.

Table 2. Log-linear model of net returns of pension funds.

	2012Q1-2023Q1		2012Q1-2017Q2		2017Q3-2023Q1	
	Coeff.	(CSE)	Coeff.	(CSE)	Coeff.	(CSE)
First panel: no scale corrections						
Constant	0.02	(1.48)	-0.22	(1.67)	0.36	(2.19)
Log(Management cost in %)	-0.48	(0.17)	-0.33	(0.18)	-0.70	(0.26)
Log(Performance fees in %)	0.33	(0.04)	0.28	(0.04)	0.43	(0.04)
Log(Volatility returns, in %)	1.28	(0.10)	1.19	(0.13)	1.41	(0.07)
R^2, adjusted	26.3		22.1		34.7	
Number of observations	7,290		4,356		2,934	
F-test on model	871		411		518	
Second panel: Management costs corrected for scale economies						
Constant	-3.48	(0.78)	-2.90	(1.08)	-4.25	(0.64)
Log(Management cost in %)*	-0.03	(0.17)	0.13	(0.17)	-0.04	(0.27)
Log(Performance fees in %)	0.30	(0.04)	0.25	(0.04)	0.38	(0.05)
Log(Volatility returns, in %)	1.31	(0.08)	1.24	(0.11)	1.41	(0.07)
R^2, adjusted	24.9		21.3		32.2	
Number of observations	7,290		4,356		2,934	
F-test on model	805		394		465	
Third panel: Management costs and performance fees corrected for scale economies						
Constant	-3.36	(0.68)	-2.65	(1.08)	-4.18	(0.65)
Log(Management cost in %)*	0.47	(0.17)	0.59	(0.16)	0.30	(0.15)
Log(Performance fees in %)*	-0.05	(0.06)	-0.21	(0.06)	0.15	(0.09)
Log(Volatility returns, in %)	1.36	(0.07)	1.26	(0.11)	1.47	(0.07)
R^2, adjusted	14.4		14.2		16.9	
Number of observations	7,290		4,356		2,934	
F-test on model	407		240		199	

Notes: CSE stands for clustered standard errors. * Refers to correction for scale.

Although economies of scale in investment cost are limited (Bikker and Meringa, 2022), the efficiency effect on profits may be related to these economies of scale. If this would be the case, that would, in principle, not change the interpretation of the observed management cost effect as measure of competition. Indeed, under competition, scale economies (may) lead to consolidation, e.g. the aircraft manufacturing industry.[4] To investigate this we correct management cost for scale economies using auxiliary regressions, see Appendix. The second panel presents the results. We see that the management effects now disappear. Apparently, the statistically significant effects of management costs on returns were mainly caused by scale economies: effects of possible efficiency differences across pension funds of the same size seem absent. We obtain the same result – i.e. no longer any effect of management costs on returns – when we estimate Eq. (5) with fixed effects for pension funds, or in first differences, two approaches which also eliminate scale economy effects.

The second explanatory variable, performance fees, has a positive effect. In contrast to what has been claimed by Broeder et al. (2019), investment services delivered by investment advice companies (in return of the performance fees) increase net profits. Possibly, this observed effect is indirect: performance fees may indicate the use of more complex and rewarding investment categories as chosen by the larger pension funds which raises net returns. In the third panel we also correct performance fees for scale economies and that approach indeed erases this effect. Hence, we do not draw any conclusion with respect to the question whether investment advice earns back its own costs.

[4] Scale may also indicate an indirect effect: larger pension funds invest more in risky assets which may contribute to higher net returns. Such effect would undermine the competition interpretation of the results. In Bikker and Meringa (2024a; 2024b) we find that – surprisingly enough – larger pension fund had over the current estimation period, and on average, no higher returns than small pension funds. In any case, our 'market share' (instead off 'returns') approach in the Section "Market share", does not suffer from such indirect effect.

Finally, we observe that taking risk is rewarded, statistically significant in any model in Table 2. This is in line with financial theory: higher volatility of returns goes with higher returns, and with what we have seen in Bikker and Meringa (2024a; 2024b), where more investments in risky asset classes resulted in higher net returns.
We split the sample into two subperiods to see how these effects may change over time. We observe that the cost variables all have larger coefficients and higher t-values in later years compared to the earlier period. This may indicate that the pension industry has become more competitive. An underlying explanation for the observed developments may be that required qualifications of board members have become more stringent over time, which has affected smaller pension funds, and stimulated mergers and acquisitions of smaller institutions.

As a first robustness test, we expand our model with cross-terms, see Eq. (6), and present the estimation results in Table 3. Such cross-terms allows for different management cost effects for varying levels of performance fees or volatility, and similarly for the other two model variables. The direct effects of Columns 1-2 are difficult to interpret, so we calculate the respective total effect in Columns 3-4 by calculating the model's first derivative to, respectively, management costs, performance fees and volatility (all in logarithms), evaluated for the average pension fund. The total effect results are fairly similar to the corresponding results in Table 2, be it that the total effects are slightly (performance fees) or substantially (management costs) stronger at similarly high levels of significance.

Table 3. Log-linear model for net returns of pension funds with cross terms (2012Q1-2023Q1)

	Model coefficients		Total effects	
	Coeff.	(CSE)	Coeff.	(CSE)
Column	1	2	3	4
Constant	-9.53	(3.29)		
Log(Management cost in %)	0.90	(0.38)	-0.79	(0.23)
Log(Performance fees in %)	0.33	(0.41)	0.34	(0.05)
Log(Volatility returns, in %)	2.26	(0.35)	1.51	(0.08)
Log Management x Log Volatility	-0.14	(0.04)		
Log Performance x Log Volatility	0.05	(0.03)		
Log Performance x Log Management	-0.07	(0.05)		
R^2, adjusted	27.6			
Number of observations	7,290			
F-test on model	463			

Note: CSE stands for clustered standard errors.

As a second robustness test, we use all observations of net returns, hence both positive and negative net returns. For each pension fund we take the largest loss and add its absolute value to each of the net returns of that fund, so that all observations are positive (except the 'largest loss' which is zero). In addition, we add the average value of the ten smallest rescaled net returns to each of the observations, so that all returns are positive. Table 4 presents the corresponding estimation results.

We see the same pattern in this extended sample of net returns: (i) management cost has a statistically significant negative coefficient, which is in line with the PCS model theory: the lower pension funds' costs, the higher their net returns. This effect becomes insignificant when corrected for scale economies as in Table 2, (ii) performance fees have a significant positive effect: investment services delivered by the corresponding investment advice companies increase net returns. This effect also becomes insignificant when corrected for scale economies, not shown here, and (iii) taking risk has been rewarded (but less significantly during 2017Q3-2023Q1).

Table 4. Log-linear model of net returns of pension funds after rescaling.

	2012Q1-2023Q1		2012Q1-2017Q2		2017Q3-2023Q1	
	Coeff.	(CSE)	Coeff.	(CSE)	Coeff.	(CSE)
First panel: no scale corrections						
Constant	11.43	(1.23)	8.57	(0.71)	15.99	(1.69)
Log(Management cost in %)	-0.65	(0.15)	-0.47	(0.08)	-0.80	(0.21)
Log(Performance fees in %)	0.35	(0.04)	0.31	(0.01)	0.41	(0.04)
Log(Volatility returns, in %)	0.42	(0.04)	0.56	(0.04)	0.10	(0.05)
R^2, adjusted	18.6		15.8		27.2	
Number of observations	10,148		5,618		4,530	
F-test on model	775		353		565	
Second panel: Management costs corrected for scale effects						
Constant	6.81	(0.46)	5.04	(0.53)	10.60	(0.49)
Log(Management cost in %)*	-0.12	(0.16)	0.03	(0.17)	-0.32	(0.21)
Log(Performance fees in %)	0.32	(0.04)	0.28	(0.05)	0.39	(0.05)
Log(Volatility returns, in %)	0.45	(0.04)	0.60	(0.05)	0.11	(0.05)
R^2, adjusted	15.7		14.2		23.1	
Number of observations	10,148		5,618		4,530	
F-test on model	630		310		453	

Note: CSE stands for clustered standard errors.

We have also run the robustness test without taking logarithms of the net returns. The results, not shown here, are fairly similar to those of Table 4.

Market shares

We apply the traditional PCS model in this section to market shares (instead of net returns): a lower marginal costs value generates a higher market share. Our market shares are based on total assets of pension funds while, in our 'model 1', costs are the sum of investment and administrative costs margins. The costs effect is large, in absolute terms, and negative (as it should be in a competitive market) and highly significant; see the upper, left side of Panel I of Table 5. Economies of scale is the main driving force, as become clear when we correct the cost margins for scale effect, see first model in the upper, left side of Panel II. Alserda et al. (2018) have demonstrated that the economies of scale in administrative costs are still very large (though declining over time), while Bikker and Meringa (2022) show that investment cost provides less economies of scale. Volatility has no effect on market shares, as expected, in contrast to the net return model.

We conclude that the market share results confirm what we have found in the net return approach, namely that scale economies in marginal costs drive the consolidation process in the pension sector. Indeed, under competition, scale economies (may) lead to consolidation. Effects of possible efficiency differences across pension funds of the same size seem absent also here.

We have various robustness tests. First, we split the administrative and investment costs margins. Administrative costs are 'waste' costs, while part of the investment costs (i.e. performance fees) are the consequence of chosen investments strategies which might be successful ("Model II"). We observe indeed that administrative costs have a highly significant negative impact on market shares, whereas investment costs have a positive effect. Therefore, we next divide investment costs into the waste part, management cost, and the 'investment' part, performance fees ("Model III"). We observe that the waste type managements costs indeed do not help to achieve greater market shares (negative coefficient) but that performance fees do contribute to growth.

Second, we re-estimate the models for the subperiods 2012Q1-2017Q2 and 2017Q3-2023Q1. All results are also found in each of these two subperiods, see Panel 1. This conclusion also holds for Panel II where cost efficiency is corrected for scale, but there the significant effects all disappear after correction.

Table 5. Log-linear model for market shares of pension funds

	2012Q1-2023Q1		2012Q1-2017Q2		2017Q3-2023Q1	
	Coeff.	(CSE)	Coeff.	(CSE)	Coeff.	(CSE)
Panel I: no scale corrections						
Model 1						
Constant	-6.54	(0.26)	-7.05	(0.18)	-6.11	(0.23)
Log(Inv. + admin. costs in %)	-1.41	(0.11)	-1.32	(0.04)	-1.55	(0.17)
Log(Volatility returns, in %)	0.06	(0.06)	-0.03	(0.04)	0.13	(0.04)
R^2, adjusted	25.1		26.5		22.0	
Number of observations	10.529		5.837		4.692	
F-test on model	1764		1056		664	
Model 2						
Constant	-8.22	(0.20)	-8.17	(0.11)	-8.16	(0.22)
Log(Investment costs in %)	0.24	(0.10)	0.25	(0.03)	0.28	(0.16)
Log(Admin. costs in %)	-1.12	(0.11)	-1.19	(0.03)	-1.05	(0.18)
Log(Volatility returns, in %)	0.02	(0.04)	0.03	(0.03)	0.04	(0.04)
R^2, adjusted	40.7		45.0		35.2	
Number of observations	10,352		5,705		4,647	
F-test on model	2365		1555		841	
Model 3						
Constant	-11.40	(1.22)	-11.74	(0.41)	-11.64	(1.74)
Log(Management costs in %)	-0.36	(0.11)	-0.32	(0.04)	-0.43	(0.17)
Log(Performance fees in %)	0.20	(0.07)	0.09	(0.02)	0.29	(0.07)
Log(Admin. costs in %)	-0.61	(0.13)	-0.56	(0.03)	-0.66	(0.15)
Log(Volatility returns, in %)	0.15	(0.05)	0.18	(0.05)	0.06	(0.05)
R^2, adjusted	32.9		28.9		37.1	
Number of observations	4,118		2,459		1,659	
F-test on model	506		251		245	
Panel II: Corrections for economies of scale						
Model 1						
Constant	-8.05	(0.51)	-9.01	(0.65)	-6.97	(0.46)
Log(Inv. + admin. costs in %)*	0.02	(0.12)	-0.16	(0.23)	0.50	(0.35)
Log(Volatility returns, in %)	-0.11	(0.12)	-0.30	(0.16)	0.10	(0.11)
R^2, adjusted	0.1		1.3		1.9	
Number of observations	10,529		5,837		4,692	
F-test on model	6		39		47	
Model 2						
Constant	-7.30	(0.53)	-7.68	(0.60)	-6.95	(0.48)
Log(Investment costs in %)*	0.12	(0.12)	0.05	(0.14)	0.27	(0.13)
Log(Admin. costs in %)*	-0.02	(0.21)	-0.29	(0.17)	0.42	(0.30)
Log(Volatility returns, in %)	0.05	(0.13)	-0.02	(0.14)	0.08	(0.12)
R^2, adjusted	0.3		0.9		3.7	
Number of observations	10,097		5,530		4,567	
F-test on model	10		17		57	
Model 3						
Constant	-7.63	(0.49)	-8.05	(0.57)	-7.35	(0.43)
Log(Management costs in %)*	0.11	(0.11)	0.02	(0.13)	0.25	(0.11)
Log(Performance fees in %)	-0.14	(0.02)	-0.14	(0.02)	-0.15	(0.02)
Log(Admin. costs in %)*	-0.21	(0.15)	-0.29	(0.16)	-0.01	(0.20)
Log(Volatility returns, in %)	0.09	(0.12)	0.03	(0.14)	0.11	(0.10)
R^2, adjusted	13.2		12.7		16.5	
Number of observations	10,060		5,512		4,549	
F-test on model	381		200		224	

Notes: CSE stands for clustered standard errors. * Refers to correction for scale.

Finally, we estimate the market share variant of Eq. (5) with fixed effects for pension funds, or in first differences, two approaches which also eliminate scale economy effects (results not shown). The administration cost margin effects on market shares, however, remain negative and statistically significant.

Conclusions

This chapter sheds light on the functioning of the pension fund market in the Netherlands. Pension funds are shielded from fierce competition by the so called Great Obligation ('grote verplichtstelling'), which ensures a mandatory participation by employees in their company's or industrial sector's own pension fund. However, some competition remains due to mergers between pension funds and take-overs of smaller pension funds, and the competition for asset managers. The Dutch supervisory authority may play an important role through regulation of pension funds, for example by setting more stringent requirements for pension board members over time.

The literature tells us that properly functioning markets are those markets in which efficiency among firms is rewarded by profitability. Therefore, we look at the relationship between efficiency and profitability of pension funds and find that this relationship is statistically significant, but the strength of the relationship is (at a β of -0.48) relatively weak, as can be derived from the low value of β compared to the results for the life insurance and banking sectors. This is in line with the fact that pension funds are shielded from fierce competition compared to these sectors. To achieve this result, we correct for the investment risk that pension funds take. Here, too, we find a positive effect of risk on profits, indicating that risk taking is rewarded. The observed efficiency-profitability relationship depends fully on scale economies in efficiency. In that sense we describe the consolidation process in the pension fund market where larger pension funds are more attractive as merger partner due to cost economies.

The relationship between efficiency and profits has strengthened over time, indicating that the increased regulation of transparency of costs has borne fruit, and may have provided incentives to reduce unnecessary costs. The increased quality of the pension fund boards and the increased professionalism of asset managers due to consolidation among pension funds may have contributed to this result. Also, competition among pension fund service providers, such as administration offices, may be reflected in the increase of the pension fund efficiency and net profit relationship.

Our overall conclusion is that in the pension fund market efficiency has been rewarded in the last decade and that this functioning has improved over time, although considerable room for improvement remains, given the low impact of efficiency on profits compared to e.g. the life insurance industry and the banking sector.

References

Alserda, G.A.G., J.A. Bikker, S.G. van der Lecq, 2018, X-efficiency and economies of scale in pension fund administration and investment, *Applied Economics* 50 (48), 5164-5188.

Bikker, J. A. 2017. Competition and scale economy effects of the Dutch 2006 healthcare insurance reform, *Geneva Papers on Risk and Insurance—Issues and Practice* 42, 53–78.

Bikker, J. A., M. van Leuvensteijn, 2008. Competition and efficiency in the Dutch life insurance industry. *Applied Economics* 40, 2063–2084.

Bikker, J.A., J.W.B. Bos, 2008, Bank Performance: A theoretical and empirical framework for the analysis of profitability, competition and efficiency, *Routledge International Studies in Money and Banking*, Routledge, Londen & New York.

Bikker, J.A., M. van Leuvensteijn, 2014, Measuring Competition in the Financial Industry: The Performance-Conduct-Structure Indicator, *Routledge*, Londen & New York.

Bikker, J.A., L. Spierdijk (eds), 2017, *Handbook of Competition in Banking and Finance*, Edward Elgar.

Bikker, J.A., J.J. Meringa, 2022, Have scale effects on cost margins of pension fund investment portfolios disappeared? *Applied Economics* 54 (39), 4501-4518.

Bikker, J.A., J.J. Meringa, 2024a, Grootste pensioenfondsen beleggen niet beter (Large pension funds do not invest better), *Economisch Statistische Berichten*, 109 (4830), 22 February, 82-85.

Bikker, J.A., J.J. Meringa, 2024b, Large pension funds do not invest more effectively than smaller pension funds, *DNB Working Paper 822*, De Nederlandsche Bank, Amsterdam.

Bikker, J.A., J.G.J. Bekooij, 2024, Market forces in healthcare insurance: the impact of healthcare reform on regulated competition revisited, *International Journal of Finance and Economics* 29 (4), 4303–4318.

Bresnahan, T., 1982. The oligopoly solution concept is identified. *Economic Letters* 10, 87–92.

Broeders, D.W.G.A., A. van Oord, D.R. Rijsbergen, 2019, Does it pay to pay performance fees? Empirical evidence from Dutch pension funds. *Journal of International Money and Finance* 93, 299–312.

Elzinga, K.G., D.E. Mills, 2011. The Lerner index of monopoly power: origins and uses. *American Economic Review*. 101 (3), 558–564.

Giocoli, N., 2012. Who invented the Lerner Index? Luigi Amoroso, the dominant firm model, and the measurement of market power. *Review of Industrial Organization* 41 (3), 181–191.

Genesove, D., W.P. Mullin (1998) Testing static oligopoly models: Conduct and cost in the sugar industry 1890 -1914, *Rand Journal of Economics* 29: 355 377.

Hay, D.A., G.S. Liu (1997). The efficiency of firms: what difference does competition make? *The Economic Journal* 107, 597–617.

Lau, L., 1982. On identifying the degree of competitiveness from industry price and output data. *Economics Letters* 10, 93–99.

Leibenstein H., 1966. Allocative Efficiency vs. "X-Efficiency", *American Economic Review* 3 (56), 392-415

Panzar, J., J. Rosse, 1982. Structure, conduct and comparative statistics. *Bell Laboratories Economic Discussion Paper*.

Panzar, J., J. Rosse, 1987. Testing for "monopoly" equilibrium. *Journal of Industrial Economics* 35 (4), 443–456.

Van Leuvensteijn, M., J.A. Bikker, A.A.R.J.M. Van Rixtel, C. Kok Sørensen, 2011. A new approach to measuring competition in the loan markets of the euro area, *Applied Economics* 43 (23), 3155–3167.

Xu, B., A. van Rixtel, M. van Leuvensteijn, 2016. Measuring bank competition under binding interest rate regulation: the case of China, *Applied Economics*, 48 (49), 4699–4718.

Appendix. Corrections for scale effects

Table A.1 presents estimates of auxiliary regressions which are used to correct for possible scale effects of the cost margins in the models for net returns and market shares (Tables 2-5). Pension fund sizes are measured by their amounts of total assets. The residuals of these auxiliary regressions have been used as corrected cost margins. Management costs have small but statistically significant scale effects of roughly -5%. Larger pension funds have lower management cost margins. This is in line with the economy of scale effects which have been found in Bikker and Meringa (2022). Performance fees do not reflect economy of scale effects: on the contrary, they increase significantly with pension fund size. Investment costs, that is the sum of management costs and performance fees, have no significant relationship with pension fund size. Scale effects in administration costs are huge with -39%, in line with the economies of scale estimates of Alserda et al. (2018). The scale effect of the total cost margin, investment costs and administration costs, is, at -18%, in between the investment (-0.02) and administration (-0.39) effects.

Table A.1. Auxiliary regressions to correct for scale effects (2012Q1–2023Q1)

	Log(Assets)	Number of observations
Table 2-4		
Log(Management costs in %)	-0.05 (0.02)	10,315
Log(Performance fees in %)	0.20 (0.05)	4,118
Table 5		
Model I		
Log(Inv. + admin. costs in %)	-0.18 (0.02)	10,529
Model II		
Log(Investment costs in %)	-0.02 (0.02)	10,352
Log(Administration costs in %)	-0.39 (0.03)	10,274

Note: Clustered standard errors are between brackets.

Monetary regime shifts

Peter Schotman

Introduction

Clemens Kool and I have known each other back from the days when we were both writing our PhD dissertation at Erasmus University in Rotterdam. Although working at different departments we shared the same supervisor and were both doing time series analysis of interest rates, inflation, and monetary aggregates. Our methods couldn't be more different, however. Clemens worked on very elaborate time-varying parameter models for which he implemented a Multi-State Kalman Filter (Kool, 1982, 1989) with great eye for detail to make the method work for interesting applications in monetary economics. My own work at the time focused on stable long-term relations following the developing cointegration literature. I was also strongly influenced by the econometric philosophy of David Hendry with its emphasis on diagnostic testing. Parameter stability was one of the important criteria for a well-designed econometric model. Controversies about the usefulness (or not) of time-varying parameter models has long since been settled. For short-term forecasting parsimonious models with adaptive parameters have been a basic tool, while time-varying vector autoregressions have become common in monetary policy analysis (Primiceri, 2005).

One of the applications of the MSKF in Clemens Kool's dissertation (later published as Kool (1995)) concerns testing for a break in the time series process of short-term interest rates in the 1910s. The economic motivation for the test is not so much whether there was a structural break, but more the timing of the break as that could shed light on the cause of the break. In various studies, referenced in Kool (1995), different causes were suggested: the founding of the Fed and interest rate targeting, war finance for WW1, suspension of the Gold Standard by Great Britain. Using data for both US and UK, Kool (1995, p 377) concludes that 'both countries started targeting interest rates shortly after their respective entry into World War I to be able to finance their war expenditures'. The MSKF dating of the break is far more convincing than earlier empirical tests had been, mainly because the MSKF accounted for outliers and heteroskedasticity.

The structural break concerns a change from a stationary first order autoregressive model to a random walk, with or without a concurrent change in the volatility. As this research was done in the late 1980s distinguishing between stationary I(0) and non-stationary I(1) time series was one of the hot topics in the widespread rise of cointegration studies with macro and monetary variables. It also largely shaped my own dissertation (Schotman, 1989).

Both Kool (1989, ch 6; 1995) and Mankiw et al. (1987) were debating breaks and regime changes before the Hamilton (1989) Markov switching model was published (and also do not refer to a possibly early working paper version of it). The Hamilton model has since become the standard tool for analyzing regime changes with more than 13,000 citations (March 2025) in Google Scholar. It is far more general than the permanent break model in Mankiw et al. (1987), but in its basic form less sophisticated than the MSKF technique developed in Kool (1989).[1] The benefit of the Hamilton model is that it is very easy to apply in many different settings. Nowadays its application is like pushing a button in many of the popular statistical software platforms. Application of the MSKF requires expert user knowledge for setting various hyperparameters, discussed at length in Kool (1989, ch 5+6).

I will use the Hamilton model to re-examine the results in Kool (1995) and Mankiw et al. (1987), where I will try to use a specification that gets as close as possible to the spirit of the MSKF. Just because the Hamilton model has become the dominant tool for analyzing regime changes, it is interesting to evaluate if it agrees on the timing of the break in the application to monetary policy in the early twentieth century. How

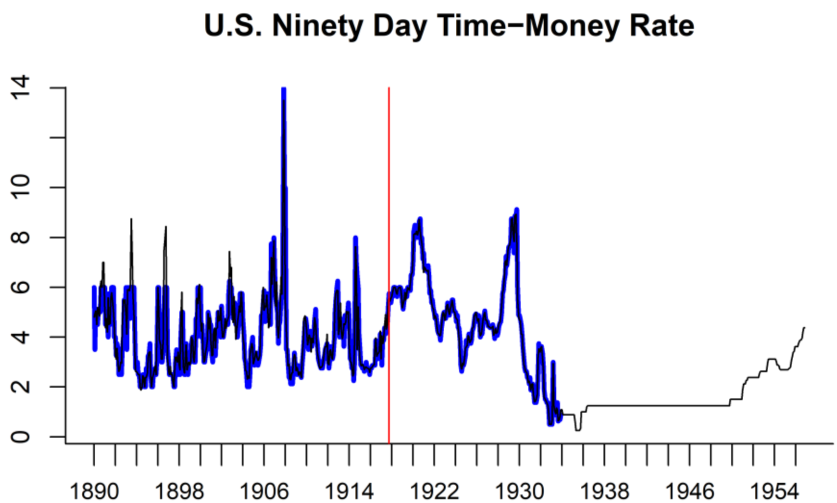
[1] Over time, however, numerous extensions of the basic Markov switching model have been proposed.

does the current standard model compare to a customized technique from the mid-1980s?

Data

The historical interest rate data series in Kool (1995) is the 90-Day Time Loan Rate at New York Banks, which is analyzed in Barsky et al. (1988) and has been published as an appendix to Mankiw et al. (1987). I will refer to it as the MMW data. It is a monthly time series for the period 1890–1933. Kool and Thornton (2004) note that the series has some measurement problems. Before 1903 it was officially capped at 6%, while outliers in late 1907 and early 1908 may well be measurement errors.

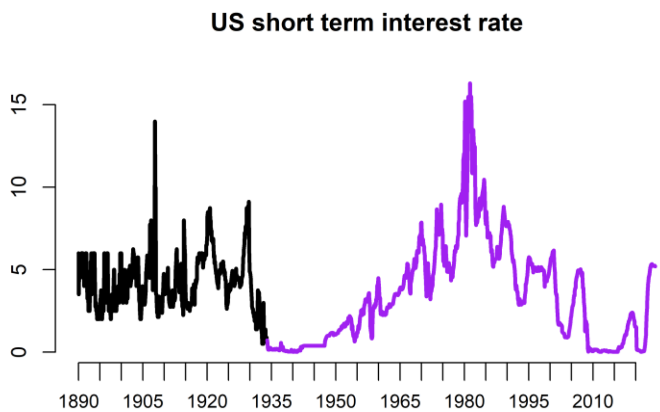
Figure 1: US short term interest rate: 1890 - 1956



Note: The figure shows the U.S. Ninety Day Time-Money Rates On Stock Exchange Loans, New York City, both from the appendix in Mankiw et al. (1987) (thick blue line) and also downloaded from NBER as series m13003 (thin black line). The red vertical line is the breakpoint identified by Kool (1995) using the MMW data.

Because of the possible data errors, I also downloaded a data series with the same name from the NBER macro history database. That data series exists for the period 01/1890- 11/1956. The NBER notes explain that the primary source of the early data is Macaulay (1938) with additional data from the Federal Reserve Bulletin. Figure 1 shows a time series plot of both data series. Until 1933, the end of the MMW data, the two series are almost identical except for the absence of the 6% cap in the very early years, and some mitigation of the outliers at the end of 1907. Kool (1995) identified a structural break in the interest time series in October 1917 (red vertical line in the graph). Without any formal modelling figure 1 suggests that the short-term interest rate became much smoother after the break, supporting a conclusion that the interest rate has become non-stationary since 1917. Still, statistically distinguishing between stationary or non-stationary behaviour of a time series is difficult with short time series. To evaluate if the interest rate really has become non-stationary since 1917, we would need to see a longer history. The extension of the same data series beyond 1933 is not very useful, however, since it is completely different from the earlier years due to the long period of interest rate pegging starting in the mid-1930s, showing up as an almost flat line until November 1949 in Figure 1.[2]

Figure 2: US short term interest rate: 1890 - 1956



Note: The figure shows the data from Mankiw et al. (1987) used in Kool (1995) (black) extended by the US 3-Month Treasury Bill rate from FRED (purple).

[2] See, however, Kool and Thornton (2004) for empirical work on the data after 1933.

Since the U.S. Ninety Day. . . New York City rate also has been discontinued, I use the 3-month TBill rate, available from the FRED database since 1934, as an alternative in order to extend the data to July 2024. Figure 2 shows the NBER data until 1933 and the 3-month TBill from 1934 onwards. The 3-month TBill rate has the similar very smooth behavior for most of the interest rate peg period. But even in later years it appears much smoother than the early data.

Interest rate regimes

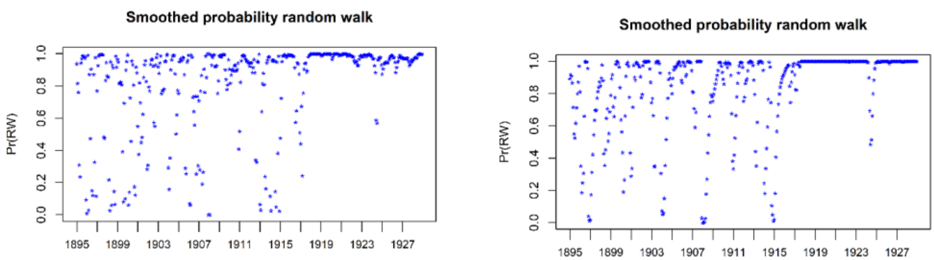
My basic model specification is identical to Kool (1995) and Mankiw et al. (1987). It is an AR(1) with time-varying parameters for the short-term interest rate R_t,

$$R_t = c_{s_t} + \rho_{s_t} R_{t-1} + \sigma_{s_t} \epsilon_t$$

Both the regression parameters c and ρ as well as the volatility parameter σ are time-varying depending on the regime indicator s_t. As in Kool (1995), I assume four different regimes characterized by four sets of parameters. The difference with Kool (1995) is in the specification of the transition between regimes. Following Hamilton (1989) I assume that the transitions are driven by a (4×4) matrix of transition probabilities with constant elements $P_{ij} = \Pr(s_{t+1} = j | s_t = i)$. Regimes 1 and 2 are both a random walk with $c = 0$ and $\rho = 1$, while regimes 3 and 4 are stationary with $c > 0$ and $0 < \rho < 1$. Within both the random walk as well as the stationary regime I distinguish between low volatility and high volatility, $\sigma_1 < \sigma_2$ and $\sigma_3 < \sigma_4$. The mixture of low and high volatility within both dynamic models allows for non-normal errors and time-varying heteroskedasticity. It can also accommodate outliers.

An important output of the Hamilton model are the smoothed regime probabilities. For each time period they are the best estimate of $\Pr(s_t = j)$, given the parameters of the model and the entire data sample. In figure 3, I show the sum of the probabilities of being in the random walk regimes 1 and 2. The results do not allow a clear timing of the breakpoints. Way before 1914 or 1917 the model sometimes assigns a probability close to one to the Random Walk (RW) model, with either high or low volatility. Later in the sample more and more time periods are identified as the RW regime. Most consistent with Kool (1995) is that after late 1917 we don't see any observation that would be classified as stationary (defined as $\Pr(\text{RW}) < 0.5$). Replacing the MMW data by the NBER data over the same sample period leads to somewhat better identification of a break. Except from a short episode in 1924, all data from 1917 onwards have a probability equal to one of being in the random walk regime.

Figure 3: Regime probabilities



Note: The graphs show the smoothed probability that the interest rate is in a random walk regime. Left panel are results for the MMW data; right panel for NBER data over the same sample period. Probabilities are based on estimated parameters in a 4-regime Markov switching model.

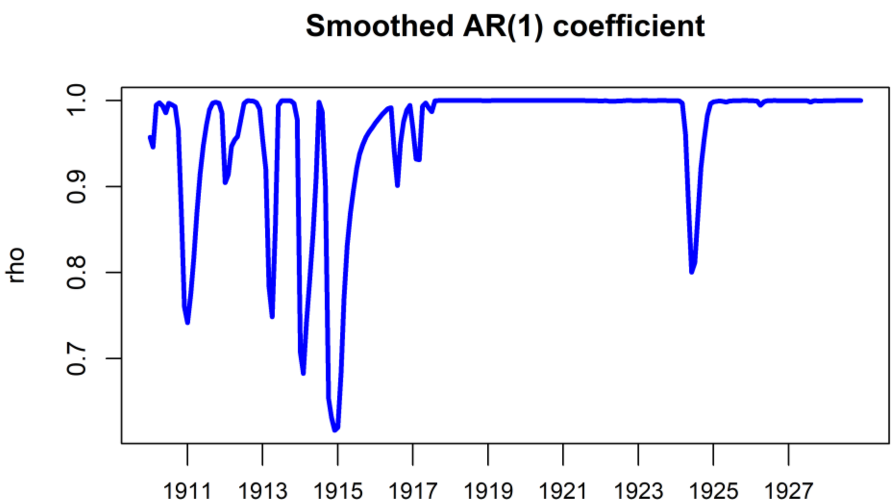
Another way to show the time-variation is by plotting the smoothed coefficient estimates E$[\rho_t]$ conditional on the model and the full data sample. The results are shown in figure 4. The AR(1) coefficient fluctuates until the end of 1917, after which is nearly identical to one for the remaining period. The plot supports the claim in Kool (1995) that the break is not as early as 1914, but more likely in 1917.

The results, especially using the NBER modified data, thus suggest that a permanent regime switch occurred just before or in 1917. However, the Hamilton model is Markovian with transition probabilities that also allow for occasional transitions back to the stationary regime. The unconditional probabilities of the four regimes are (using NBER data):

	RW low vol	RW high vol	AR(1) low vol	AR(1) high vol
Probability	0.59	0.24	0.12	0.05

The system will be in a RW regime 73% of the time. The break is thus not permanent. The model implies that none of the regimes is an absorbing state, meaning that when it is entered, the process will stay there forever. That may be a defect of the Markov switching model, even though in theory the transition probabilities could be such that once in the low volatility RW regime, the interest rate will remain there forever. Kool's MSKF model also allows for a return to stationarity if the data would imply it. The model in Mankiw et al. (1987) does not.

Figure 4: AR(1) Parameter Estimates



Note: The figure shows the smoothed estimates of the AR(1) parameter ρ_{s_t} for the NBER data. As in the corresponding figure 3 in Kool (1995), the plot zooms in on the period 1910–1928.

Extending the data to 2024, one would expect many different regimes, including the interest rate peg, the Volcker era, and the recent zero lower bound. Nevertheless, estimating the same model on this long data set, leads to the surprising result that the stationary regime ceases to exist. The model only strongly distinguishes between different volatilities. The Maximum Likelihood estimates set the AR(1) parameter ρ to a value above 0.99, effectively deciding that within the class of 4-regime Hamilton models a random walk with time-varying volatility is the best description of the data over the entire 135 years sample period. As the AR and RW regimes are so similar, the model is nearly unidentified and the transition probabilities can not be precisely estimated.[3] So the shift to a random walk in the early twentieth century may have been permanent after all. To enforce the existence of a stationary regime I ran the model again with the restriction that $\rho < 0.97$ in the stationary regimes. The fit of this model is, however, substantially worse: more than 20 points difference in log likelihood for the same number of parameters. Ending up at the boundary of all random walks does not appear a local optimum. Evidence for stationarity in 135 years of interest rate data is very limited.[4]

Time-varying persistence and the term structure

The issue of time-varying parameters and transitions from and to random walk and stationary regimes has wider applications than the interpretation of historic events more than a century ago. At the time Clemens and I were working on our PhD, Campbell and Shiller (1987) showed that when the short-term interest rate has a unit root, standard expectations models of the term structure can account for much of the volatility of long-term rates. In a similar vein, Mankiw and Miron (1986) showed that the spread between long and short- term rates has different predictive power depending on whether the short-term rate is a random walk or stationary. The idea is

---

[3] I estimate my model in R as this allows me to restrict the parameters in the RW regime to (0, 1). Unrestricted estimation in EViews for a 2-regime Markov switching model does give me a stationary regime, but at the cost of an explosive model ($c > 0, \rho < 1$) in the non-stationary regime.
[4] A serious time series model for the very long data series is outside the scope of this note. It would need more regimes and more than AR(1) dynamics.

that if the short-term interest rate is a random walk, changes in interest rates are not predictable and all shocks are permanent. Hence long-term rates should also be (near) random walks. On the other hand, if interest rate shocks are transitory, the spread between long- and short-term interest rates should carry information about expected future changes in the level of interest rates. If the spread is currently high, then the market apparently expects a future increase in the level of interest rates. Kool and Thornton (2004) use this classic insight for term structure models to obtain indirect evidence of a regime change.

Over time we have learned that neither I(0) nor I(1) models can explain the volatility of long-term interest rates. We need something in between and time-varying parameter models can generate such dynamics. A regime-switching model, where most of the time the short-term interest rate behaves as a random walk, and sometimes it reverts to 'normal' levels, can explain the yield curve much better. Ideas from the early work of Clemens Kool in his dissertation have always been on the top of my mind. In my own work we used this idea in Pfann et al. (1996), where we proposed an AR(1) with time-varying parameters to explain the term structure. Different from the Hamilton model we modelled the transitions as a function of the level of interest rates themselves.

Conclusion

We often evaluate published models on fresh out-of-sample data. In this paper I check if the results of an old, published paper are robust to newer methods that did not exist at the time. Methodology in the original was customized to the problem and required expert knowledge, which could have resulted in overfitting. The newer off-the-shelf method is a very standard tool that hardly requires any tuning or expert knowledge. The comforting conclusion from the exercise is that the results of the original Kool (1995) paper are robust to this change in methods. And, if anything, the results are strengthened by the modified data available in the NBER historical archives. My conclusion is not entirely new, since Newbold et al. (2001) earlier came to the same conclusion using Smooth Transition models, another popular technique to model time-varying parameters.

References

Barsky, R., N. Mankiw, J. Miron and D. Weill (1988). The Worldwide Change in the Behavior of Interest Rates and Prices in 1914. *European Economic Review 32*, 1123–1154.

Campbell, J., and R. Shiller (1987). Cointegration and Tests of Present Value Relations. *Journal of Political Economy 95*, 1062–1088.

Hamilton, J., (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica 57*, 357–384.

Kool, C., (1982) The Multi-State Kalman Filter. *Carnegie Rochester Conference Series on Public Policy 17*, 39–46.

——— (1989). *Recursive Bayesian Forecasting in Economics: The Multi State Kalman Filter Method*. PhD dissertation, Erasmus University.

——— (1995). War Finance and Interest Rate Targeting: Regime Changes in 1914-1918. *Explorations in Economic History 32*, 365–382.

Kool, C., and D. Thornton (2004) "A Note on the Expectations Hypothesis at the Founding of the Fed", *Journal of Banking and Finance 28*, 3055–3068.

Macaulay, F., (1938) "Appendix A: Tables," In *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1856*, Boston: NBER.

Mankiw, N., and J. Miron (1986) "The Changing Behavior of the Term Structure of Interest Rates", *Quarterly Journal of Economics 101*, 211–228.

Mankiw, N., J. Miron and D. Weill (1987) "The Adjustment of Expectations to a Change in Regime: A Study of the Founding of the Federal Reserve", *American Economic Review 77*, 358–374.

Newbold, P., S. Leybourne, R. Sollis and M. Wohar (2001) "U.S. and U.K. Interest Rates, 1890-1934: New Evidence on Structural Breaks," *Journal of Money Credit and Banking 33*, 235–250.

Pfann, G., P. Schotman, and R. Tschernig (1996) "Nonlinear Interest Rate Dynamics and Implications for the Term Structure", *Journal of Econometrics 73*, 149–176.

Primiceri, G., (2005) "Time Varying Structural Vector Autoregresssions and Monetary Policy", *Review of Economic Studies 72*, 821–852.

Schotman, P., (1989). *Empirical Studies on the Behavior of Interest Rates and Exchange Rates*. PhD Dissertation, Erasmus University.

Convergence vs. Divergence in Emerging Market Sovereign Spreads

Bertrand Candelon and Metodij Hadzi-Vaskov

Introduction

Sovereign spreads in emerging market economies (EMEs) are typically found to be driven by both domestic or country-specific and global factors. In this context, the movements of sovereign spreads over time can be generally viewed as a balancing act between external (global) factors leading to common dynamics that give rise to convergence across emerging market sovereign spreads and country- or region-specific characteristics implying divergence across EMEs. Using a compact VAR setting, this study provides evidence on the implications of the most recent shock episode—the onset of the Covid-19 pandemic and the subsequent inflation surge (2021-23)—for sovereign spreads in three groups of EMEs: Latin America, Emerging Europe, and Emerging Asia. In particular, the analysis aims to explore (i) the extent to which sovereign spreads showed signs of convergence or divergence both within and across EME regions, and (ii) what role the onset of the Covid-pandemic played in terms of changing these dynamics.

Determinants of sovereign spreads

The determinants of sovereign spreads in emerging market economies have been a fertile ground for academic research and policy discussions. Starting with the seminar work by Edwards (1984), many studies have attempted to shed light on the factors that have an impact on sovereign spreads in emerging market economies. Some of these studies include Ciarlone, Piselli, and Trebeschi (2009), Bellas, Papaioannou, and Petrova (2010), Baldacci, Supta, and Mati (2011), Csonto and Ivaschenko (2013), and Hadzi-Vaskov and Ricci (2022). Among others, these analyses identify a multitude of factors that play an important role in determining sovereign spreads, including various indicators capturing the fiscal situation and developments (such as gross or net public debt, fiscal balances), general macroeconomic conditions (real GDP growth, inflation), developments in the external sector (reserves, current account movements, and external debt metrics), as well as indicators that gauge global financial market conditions (VIX and U.S. interest rates).

Origins of divergence

In general, the forces driving emerging market spreads may be classified in two groups – those that lead to common movements across different EMEs and others that lead to differences among individual EMEs. On the one hand, global developments and financial conditions affect overall movement of capital flows into or out of EMEs, treating all or most EMEs as a compact group. For instance, Csonto and Ivaschenko (2013) report that global factors are the main determinants of spreads in the short run. Similarly, Bellas, Papaioannou, and Petrova (2010) report that overall financial volatility is a more important determinant of spreads than fundamentals in the short run. To some extent, this dimension is related to the literature on contagion in which change in financial conditions in one market is quickly translated into other markets, thereby amplifying the initial shock.

On the other hand, country-specific fundamentals, especially captured by changes in fiscal indicators, lead investors to distinguish between different EMEs based on their perceived risk profiles. In turn, such country- or region-specific characteristics may result in divergence of spreads among EMEs even when they are faced with the same type of global shock or global (financial) conditions. In this case, investors may critically distinguish between EMEs in search of those that are perceived as less risky (reflecting developments in fiscal, financial, or external indicators, for example), thereby implying behavior similar to "flight-to-quality" (Beber et al., 2009).

Recent shock episodes

Emerging markets seem to have shown higher resilience to global interest rate volatility in the most recent episodes compared to previous ones (GFSR, October 2023). For instance, the October 2023 GFSR finds that the average sensitivity of sovereign yields in emerging markets to US interest rates declined considerably (between two-thirds and two-fifths) in the course of the monetary policy tightening cycle in response to the inflation surge compared to the previous shock episodes, such as the taper tantrum. Nonetheless, GFSR October 2024 also points out that emerging markets and frontier economies with weak and worsening fiscal buffers have seen relative widening of sovereign bond spreads compared to other jurisdictions.

The analysis sheds light on the behavior of sovereign spreads in two period. First, it delves into the dynamics of emerging market sovereign spreads in the period from the early 2000s—including turmoil episodes such the Global Financial Crisis and the taper tantrum—until the onset of the Covid-pandemic. Second, the analysis explores the impact of several recent shocks on the convergence vs divergence of sovereign spreads in EMEs, such as: (i) the Covid-19 pandemic; and (ii) inflation surge (2021-23) and ramifications from the associated policy tightening. To our knowledge, this is the first study looking at divergence vs convergence of EM sovereign spreads around these most recent shocks.

The rest of the paper is organized as follows. Section 2 describes the dataset, and Section 3 presents the methodology and empirical strategy. Results are discussed in Section 4. Section 5 provides concluding remarks.

Data

Series on sovereign bond spreads for emerging market economies come from JP Morgan's Emerging Market Bond Index Global (EMBIG). Overall, the dataset includes sovereign spreads for six economies from Latin America (Brazil, Chile, Colombia, Mexico, Peru, and Uruguay), four economies from Emerging Europe (Hungary, Poland, Serbia, and Turkey), and four economies from Emerging Asia (China, Indonesia, Malaysia, and Vietnam). For systemic economies, we use data series for yields on 10-year U.S. Treasury bonds, 10-year German Bunds and 10-year Japanese bonds. The analysis is based on monthly observations over two time periods: pre-Covid pandemic

([January 2000 – March 2020]) and post-Covid pandemic period ([April 2020 – February 2025]).

Methodology and empirical strategy

To evaluate the divergence/convergence between the countries, we rely on traditional stationary VARX models, with the following representation:

$$y_t = c + \sum_{i=1}^{p} A_i \, y_{t-i} + B_i x_{t-i} + \varepsilon_t,$$

where y_t consists of stationary sovereign bond spreads, x_t are a set of exogenous variables composed by the 10-year sovereign bonds of the systemic countries (US, Japan, Germany and China), p the optimal lag order determined by the AIC information criterion and ε_t white noise errors with the following i.i.d ness properties E$(\varepsilon_t,) = 0$, E$(\varepsilon'_t, \varepsilon_t) = \Sigma$, and E$(\varepsilon_t, \varepsilon'_s) = 0$ for t ≠ s. Several VARXs are considered within the different regions (Latin America, Emerging Europe and Asia) but also VARX at a regional level.

Following Candelon, Luisi and Roccazzella (2022), the degree of convergence is evaluated via a fragmentation matrix, computed as the correlation of the forecasts generated by the VARX at a 6-month horizon. This measure of convergence presents the advantage of being forward-looking and not only based on past observations, but potentially biased by rare and temporary historical events.

We also complete the previous measure of convergence, by calculating the fragmentation matrix after a shock originating from US and Germany. To this aim, impulse response functions are calculated within the VARXs. The shocks are identified via a Cholesky decomposition and the country of origin of the shock is ordered first.

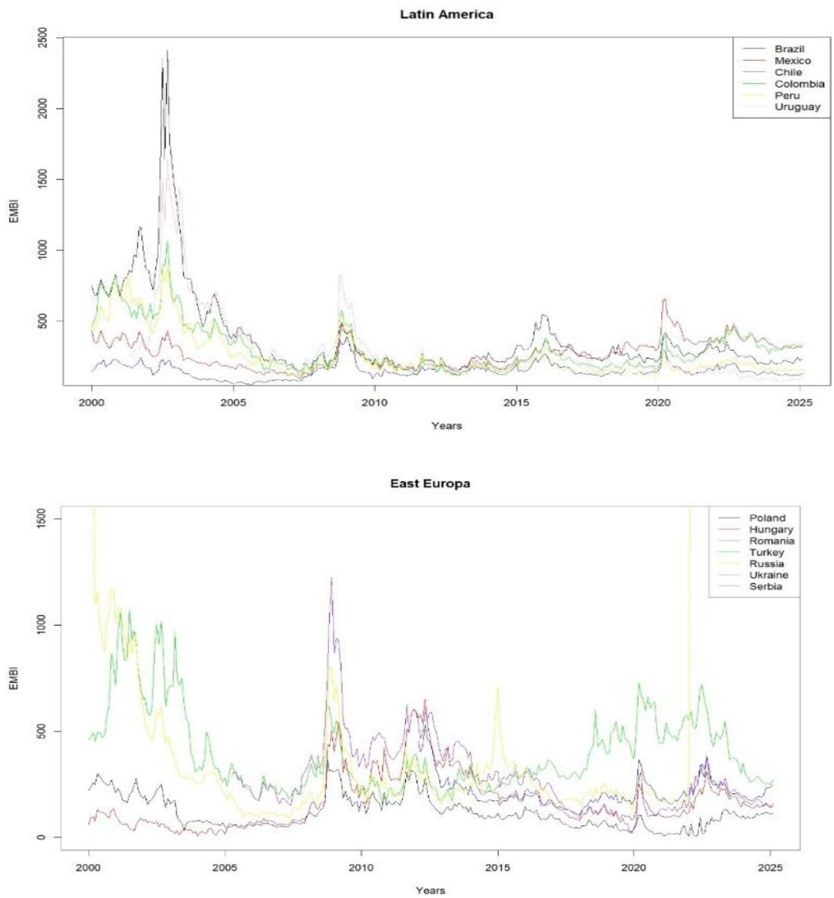
Results

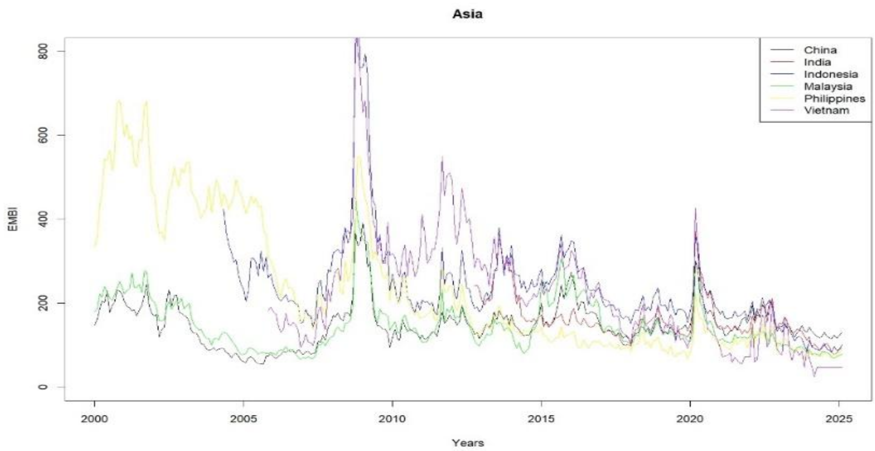
Descriptive evidence

Sovereign spreads of emerging market economies in different regions across the world have shown strong co-movement over the past quarter of a century (Figure 1).

In particular, this co-movement is particularly visible for sovereign spreads in Latin America across the full period (Figure 1, Panel A), while it has typically been gaining prominence during turmoil episodes in Emerging Europe and Emerging Asia (Figure 1, Panels B and C). For instance, the co-movements in these regions seem to have strengthened considerably during the Global Financial Crisis and the outbreak of the Covid-19 pandemic. On the other hand, Latin American economies seem to show close co-movement across the entire period, which peaked at the time of Argentina financial crisis in 2001-02.

Figure 1: Sovereign Spreads in Emerging Market Economies





Co-movement results: history, forecast and global shocks

The formal analysis reports three sets of results about the co-movement of sovereign spreads in emerging markets. The first set of results is based on historical correlations over the pre-Covid and post-Covid periods. The second set refers to results based on forecasts about correlations 12 months ahead. Finally, the third set shows results about the co-movements of sovereign spreads in emerging markets following global shocks in financing conditions, proxied by shocks to yields of 10-year U.S. Treasury bonds and yields of 10-year German Bunds.

Exploring co-movements in historical data

The results from the formal analysis, presented in Figure 2, support the descriptive evidence about the movement of sovereign spreads for the emerging economies of Latin America. In the pre-Covid period, the correlation in sovereign spreads for each bilateral country pair was at least 0.5, and in many cases well above 0.8. In particular, the co-movement in spreads has been particularly strong among Chile, Colombia, Peru, and Uruguay. Overall, the weakest co-movement in the region was observed for Mexico-Peru and Mexico-Uruguay country pairs.

Turning to the post-Covid period, the co-movements in sovereign spreads have generally remained strong across the region, with a couple of exceptions. The most visible is the decline in the correlation of Colombia's sovereign spread with the rest of the region. In particular, its correlation with Brazil, Chile, Mexico, and Uruguay dropped markedly, approaching zero in the last case. The correlations of Peru weakened somewhat too, albeit to a much lesser extent with the exception of the case of Uruguay.

Figure 2: Co-movements in Latin America

Pre-Covid

Post-Covid

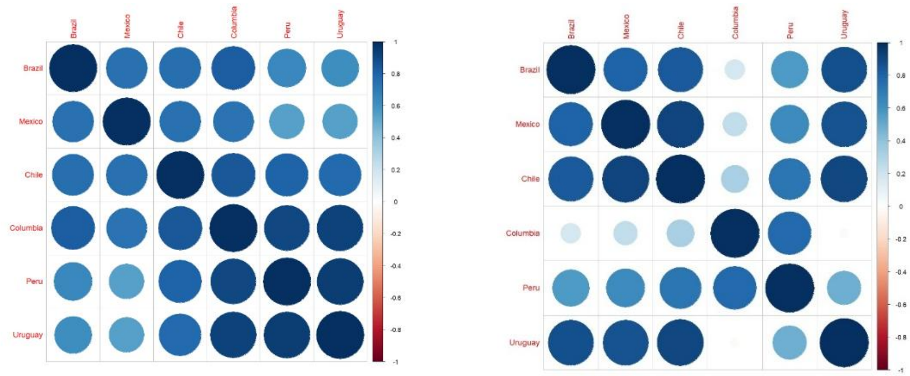
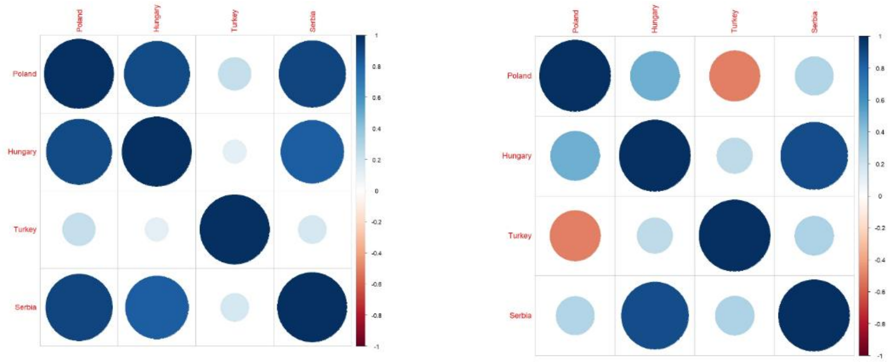


Figure 3 presents similar results from the formal analysis for the economies of Emerging Europe. Overall, the correlations of sovereign spreads have been very strong in the pre-Covid period, especially among Hungary, Poland, and Serbia, typically exceeding 0.8. In contrast, the sovereign spreads of Turkey showed much weaker co-movements with the other economies in Emerging Europe, with typical correlations in the range 0.1-0.3, largely reflecting the turbulent period following the financial crisis in Turkey in 2001.

In the period after the outbreak of the Covid pandemic, sovereign spread co-movements among the economies of Emerging Europe generally weakened, with the exception of the Hungary-Serbia country pair, where the co-movement was somewhat stronger relative to the pre-Covid period, with a correlation of above 0.8. In most other cases, correlations dropped below 0.5, In the case of the Turkey-Poland country pair, they even turned negative, implying a divergence of their sovereign spread trajectories.

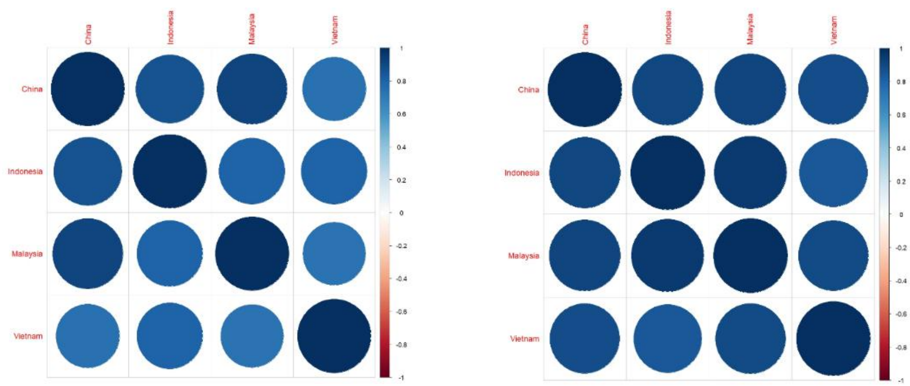
Figure 3: Co-movements in Emerging Europe
Pre-Covid Post-Covid



Turning to Emerging Asia, Figure 4 shows that co-movements of sovereign spreads were quite strong among China, Indonesia, Malaysia, and Vietnam in the pre-Covid period. For most country pairs, the correlations were well above 0.7, with the strongest being between the sovereign spreads of China and Malaysia.

In contrast to Latin America and Emerging Europe, the co-movements of sovereign spreads among the economies of Emerging Asia further strengthened in the period after the outbreak of the Covid pandemic. Overall, these results point at considerable convergence across the entire region in the post-Covid period, with correlations exceeding 0.8 for all country pairs in our sample.

Figure 4: Co-movements in Emerging Asia
Pre-Covid Post-Covid

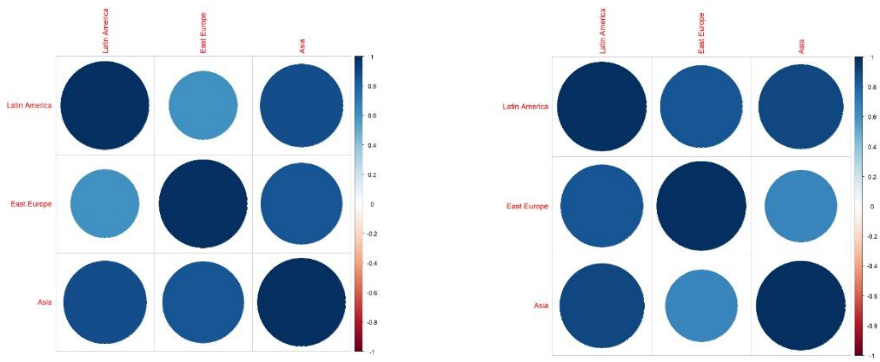


While Figures 2, 3, and 4 depicted the co-movements of sovereign spreads for individual countries within a specific region, Figure 5 turns to the co-movements among the regions themselves. The left panel shows the correlations among the average spreads for the three regions of interest in the pre-Covid period. Overall, the correlations seem quite strong, with the weakest—between Emerging Europe and Latin America—being well over 0.5.

In turn, the right panel shows the same matrix of correlations for the period after the outbreak of Covid. The general conclusion is that correlations have been strong and broadly similar to the pre-Covid period. A key observation worth underlining is that the weakest link is not Emerging Europe-Latin America but rather Emerging Europe-Emerging Asia. The co-movement of Latin America with the remaining two regions seems very similar.

Figure 5: Co-movements Across Regions

Pre-Covid | Post-Covid



Co-movements in forecasts

Correlations across 12-month-ahead forecasts for sovereign spreads of Latin American emerging economies are presented in Figure 6. Similar as the case for actual historical data over the pre-pandemic period, the left panel shows that the co-movements across forecasts of sovereign spreads for all economies of Latin America are very strong. In particular, the correlations across forecasts are even higher than for historical series, typically exceeding 0.9.

Turning to the post-Covid period, the right panel shows that the co-movements in 12-month-ahead forecasts for sovereign spreads have been generally weaker than in the pre-Covid period. In fact, for most country pairs, they weakened considerably, with the largest declines in correlations registered for country pairs involving Uruguay, and to a lesser extent, for country pairs involving Brazil. In turn, the largest decline is seen for Brazil-Uruguay, which is the only one that turns negative.

Figure 6: Co-movements Among Forecasts in Latin America

Pre-Covid Post-Covid

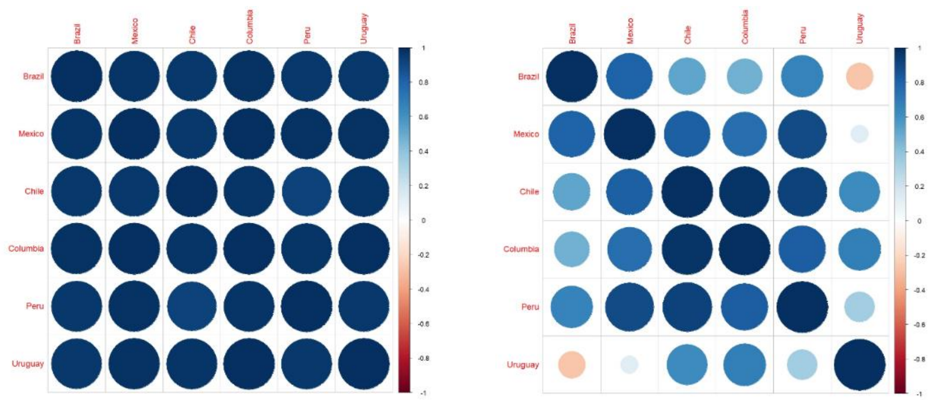
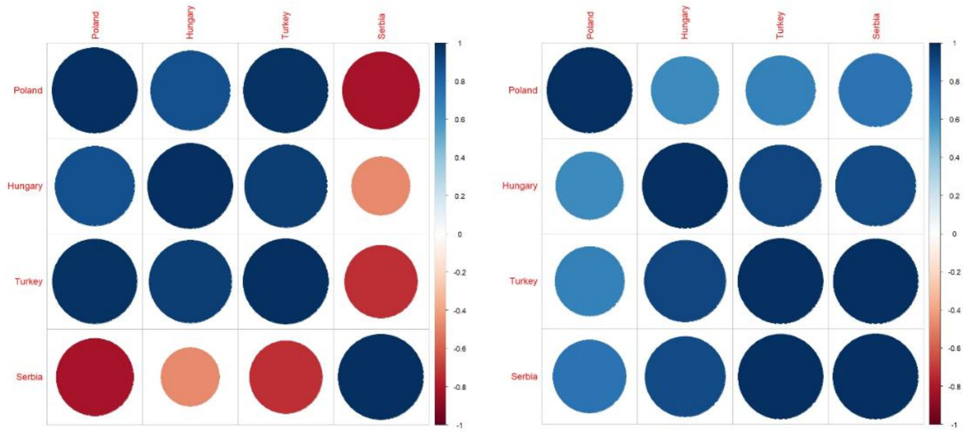


Figure 7 shows results about co-movements of 12-month-ahead forecasts for sovereign spreads of economies from Emerging Europe. These co-movements are generally found to be strong for the pre-Covid period, with correlations above 0.8 among Hungary, Poland, and Turkey. However, forecasts for Serbia's spreads seem to be disconnected from the rest of Emerging Europe, with correlations falling in the negative territory of all its bilateral pairs.

Turning to the period after the outbreak of the Covid pandemic, the right panel of Figure 7 shows that co-movement of forecasts for sovereign spreads strengthened markedly between Serbia and the rest of Emerging Europe, while it generally weakened for Poland, with the rest remaining broadly similar to the pre-Covid period.

Figure 7: Co-movements Among Forecasts in Emerging Europe

Pre-Covid Post-Covid

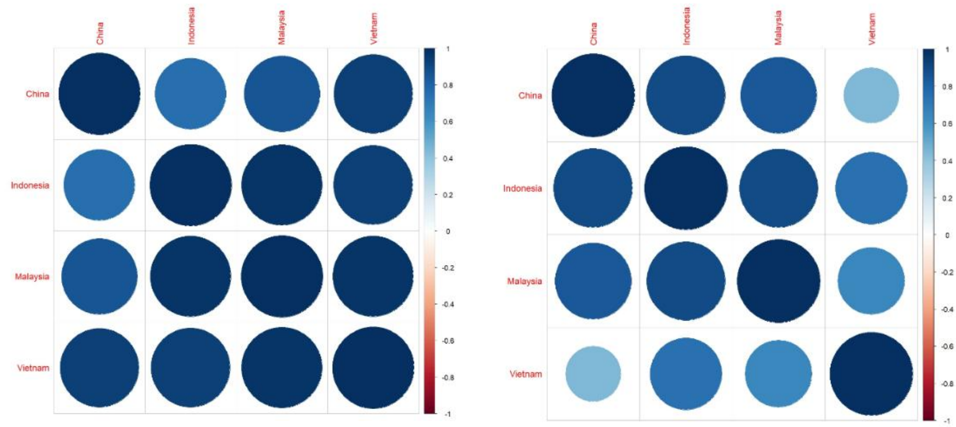


Turning to Emerging Asia, Figure 8 shows that co-movements of the sovereign spread forecasts have been very strong among China, Indonesia, Malaysia, and Vietnam in the pre-Covid period. With the exception of the China-Indonesia country pair, the other correlations typically exceeded 0.8. These results are very close to the findings for the co-movements among these economies based on the historical data for sovereign spreads.

In addition, the right panel of Figure 8 shows that the co-movements of sovereign spreads among the economies of Emerging Asia only marginally weakened following the outbreak of the Covid pandemic. Overall, the correlations among forecasts for China, Indonesia, and Malaysia remained strong, typically exceeding 0.7. In the case of Vietnam, the correlations of its forecasts with the rest of the region weakened more markedly, albeit remaining in the 0.3-0.5 range.

Figure 8: Co-movements Among Forecasts in Emerging Asia

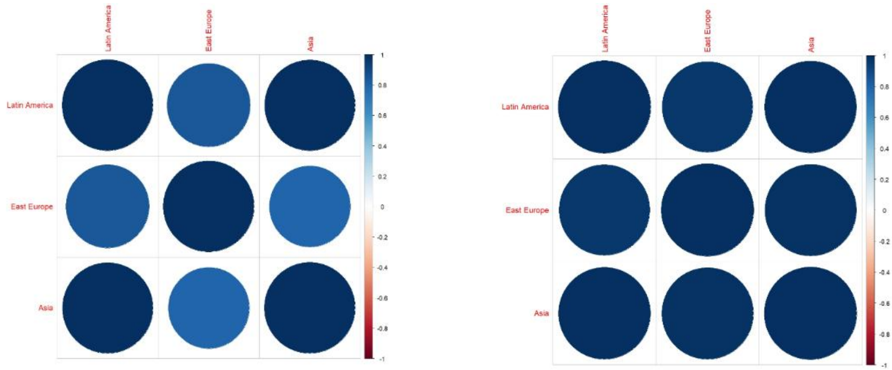
Pre-Covid Post-Covid



While Figures 6, 7, and 8 depicted the co-movements of forecasts for individual countries' sovereign spreads, Figure 9 reports co-movements among sovereign spread forecasts for the three emerging market regions. The left panel depicts the correlations among the average sovereign spread forecasts for the three regions of interest in the pre-Covid period. Overall, the correlations seem strong, typically above 0.7. Somewhat weaker correlations are reported between Emerging Europe and Latin America—similar as was documented in the case with the historical data—as well as for between Emerging Europe and Emerging Asia. Turning to the post-Covid period, the right panel shows that all correlations among sovereign spread forecasts strengthened markedly, implying remarkably convergence of forecasts for the three major emerging market regions.

Figure 9: Co-movements Among Forecasts Across Regions

Pre-Covid Post-Covid



Response to external shocks

In the period before the outbreak of the Covid pandemic, the response of sovereign spreads to a shock in U.S. Treasury yields was quite synchronous across economies in Latin America. In fact, most of the correlations among the spreads of these economies is estimated at about 0.9, with the sole exception of Chile, which response following a shock to U.S. Treasury yields was somewhat muted, albeit with a correlation vis-à-vis the other Latin American economies still above 0.5, as depicted in the top left panel of Figure 10.

The results from the formal analysis, presented in Figure 2, support the descriptive evidence about the movement of sovereign spreads for the emerging economies of Latin America. In the pre-Covid period, the correlation in sovereign spreads for each bilateral country pair was at least 0.5, and in many cases well above 0.8. In particular, the co-movement in spreads has been particularly strong among Chile, Colombia, Peru, and Uruguay. Overall, the weakest co-movement in the region was observed for Mexico-Peru and Mexico-Uruguay country pairs. The response of sovereign spread co-movements across Latin America to shocks in yields of German Bunds was weaker compared to the case of U.S. Treasuries, though still generally positive, except for a few country pairs involving Peru and Uruguay.

Turning to the post-Covid period, most co-movements in sovereign spreads in response to global shocks to the U.S. and Germany weakened considerably. This seems to be especially the case with co-movements in response to U.S. Treasury shocks, when several co-movements across the region turned negative. most visible is the decline in the correlation of Colombia's sovereign spread with the rest of the region. On the other hand, the response to shocks to German Bunds resulted in somewhat weaker co-movements across the region, albeit without divergent trajectories as implied in the case of the U.S.

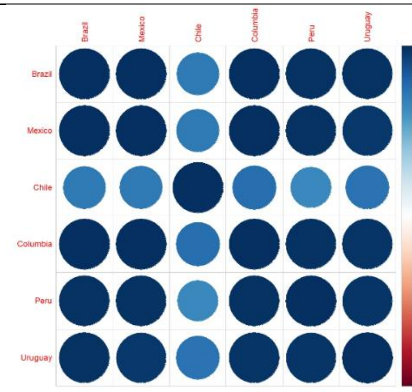
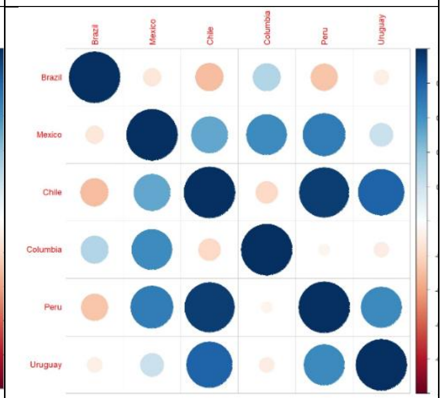
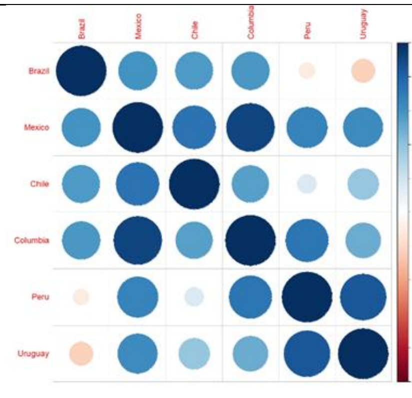
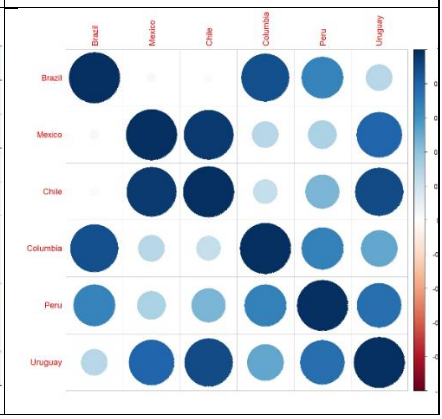
Figure 10: Co-movements in Latin America in Response to Global Shocks		
	Pre-Covid	Post-Covid
Shock to 10-year U.S. Treasuries		
Shock to 10-year German Bunds		

Figure 11 shows co-movements of sovereign spreads among the economies of Emerging Europe in response to shocks to U.S. Treasuries and German Bunds. Overall, in the period before the pandemic, the response across the region was much more synchronous in response to shocks in German Bunds (with implied correlations across the region typically around 0.9) than the response to shocks in U.S. Treasuries (positive in all cases, but with average correlations of about 0.5). The post-Covid period saw a dramatic decline in the co-movements across Emerging Europe economies in response to German yield shocks. On the other hand, the co-movements in response to U.S. shocks seem to have remained broadly similar.

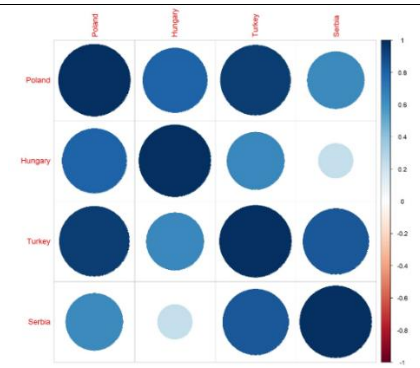
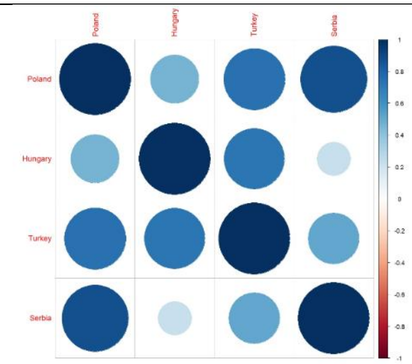
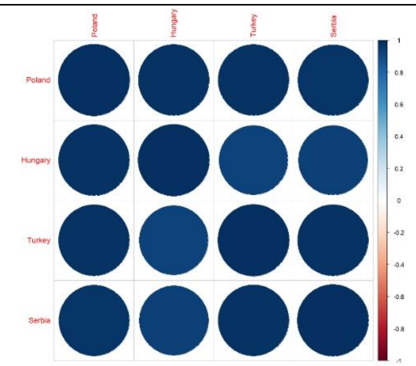
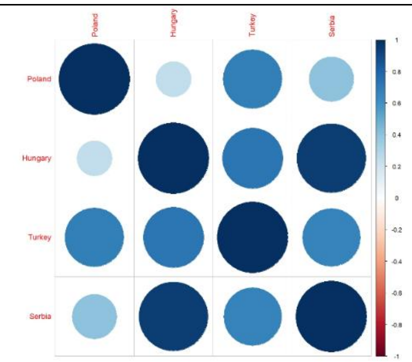
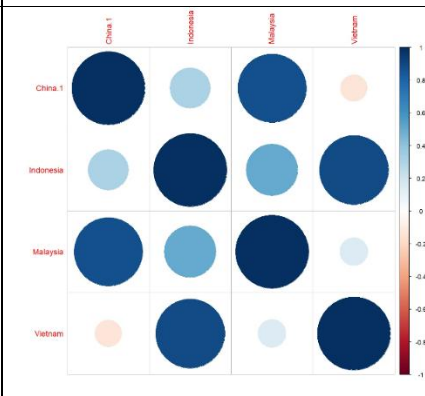
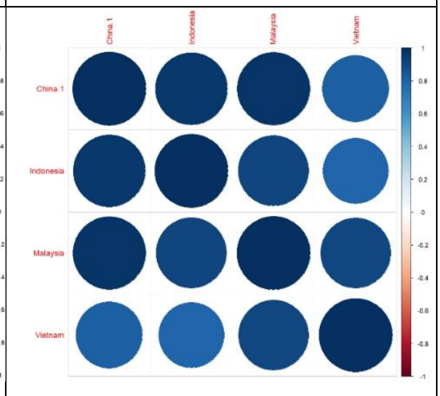
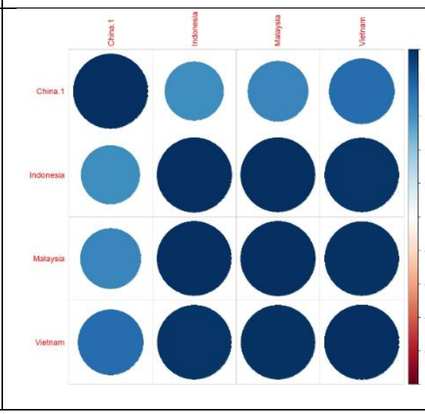
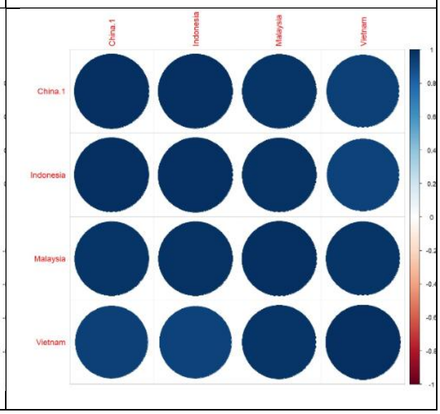
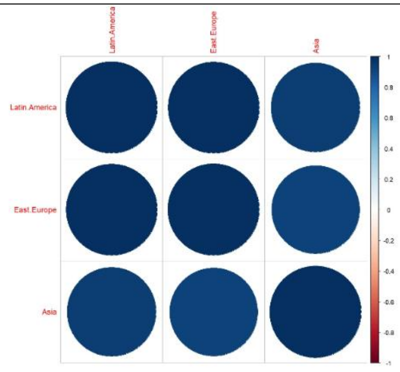
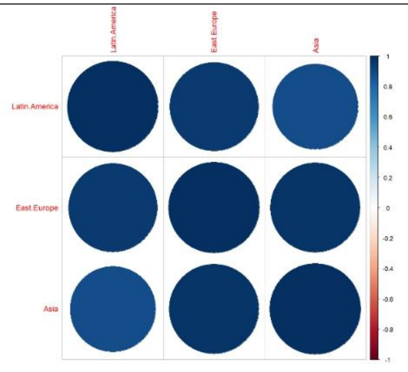
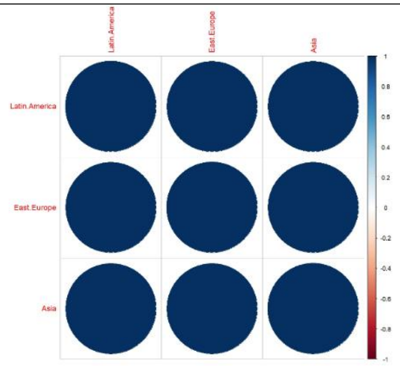
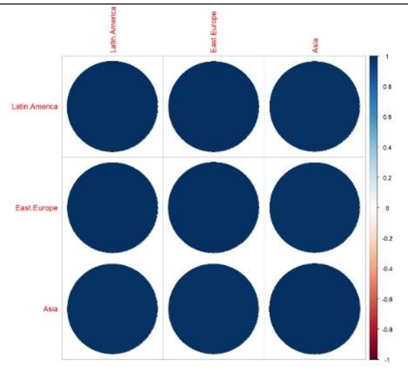
Figure 11: Co-movements in Emerging Europe in Response to Global Shocks		
	Pre-Covid	Post-Covid
Shock to 10-year U.S. Treasuries		
Shock to 10-year German Bunds		

Figure 12 shows co-movements of sovereign spreads among the economies in Emerging Asia following shocks to U.S. Treasury and German Bund yields. In the pre-Covid period, the co-movements among China, Indonesia, Malaysia, and Vietnam were considerably stronger following a shock to German than U.S. yields, with the only exception being the country pair China-Malaysia. In the period following the outbreak of the pandemic, the co-movements across this region strengthened in response to shocks to both systemic economies, a finding that stands in sharp contrast to the results for the economies within the previous two regions – while the onset of the pandemic seems to have led to divergence among the spreads within Latin America and Emerging Europe, it seems to have contributed to more convergence among the spreads of economies with Emerging Asia.

Figure 12: Co-movements in Emerging Asia in Response to Global Shocks		
	Pre-Covid	Post-Covid
Shock to 10-year U.S. Treasuries		
Shock to 10-year German Bunds		

Turning from the findings for individual economies within specific regions to co-movements of average spreads among the three regions, Figure 13 shows that the co-movements among the regions to both shocks to U.S. Treasuries and German Bunds were very strong in the pre-pandemic period. Most correlations were above 0.9 over this time period. In addition, the onset of the pandemic does not seem to have materially affected the co-movements among the regions – overall, even if co-movements of spreads changed considerably within the regions, they seem to have remained almost intact among the three regions themselves.

Figure 13: Co-movements Across Regions in Response to Global Shocks

	Pre-Covid	Post-Covid
Shock to 10-year U.S. Treasuries		
Shock to 10-year German Bunds		

Concluding remarks

Exploring the extent to which sovereign spreads showed signs of convergence or divergence both within and across the three major emerging market regions before as well as after the onset of the Covid-pandemic, our analysis arrives at four main findings.

First, the co-movements implied by actual historical data imply that sovereign spreads within the three major emerging market regions showed a high degree of convergence over the pre-pandemic period, which has generally weakened for Latin

America and Emerging Europe in the post-pandemic period, while it remained broadly similar in the case of Emerging Asia.

Second, looking at the co-movements implied by forecasts of sovereign spreads suggests that the convergence generally weakened among emerging economies within Latin America (consistent with the result based on historical series), and to a lesser extend for Emerging Asia, while the results for Emerging Europe show stronger convergence for some and weaker for other country pairs.

Third, the exercise based on the response of emerging market spreads to global shocks in systemic economies suggests that the convergence among sovereign spreads within Latin America and Emerging Europe generally weakened after global shocks affecting U.S. Treasury yields or German Bunds, or even turned from convergence into divergence for some country pairs in Latin America following a shock to U.S. Treasury yields in the period after the onset of the Covid pandemic. On the contrary, the convergence among sovereign spreads for Emerging Asia economies strengthened somewhat following shocks to yields of either U.S. Treasuries or German Bunds in the post-Covid period.

Finally, the co-movement across the three regions seems quite resilient to the different shocks, showing strong convergence both in the pre-Covid as well as in the post-Covid period. Overall, these findings that sovereign spreads across the different emerging market regions show a high degree of convergence, which—notwithstanding some noise for specific country pairs within the regions—display robust convergent trajectories that seem to have been immune to the series of recent shocks, including the onset of the Covid-pandemic.

References

Baldacci, E., Gupta, S., & Mati, A. (2011). Political and fiscal risk determinants of sovereign spreads in emerging markets. *Review of Development Economics, 15*(2), 251–263.

Beber, A., Brandt, M. W., & Kavajecz, K. A. (2009). Flight-to-quality or flight-to-liquidity? Evidence from the euro-area bond market. *The Review of Financial Studies, 22*(3), 925–957.

Bellas, D., Papaioannou, M. G., & Petrova, I. (2010). Determinants of emerging market sovereign bond spreads: Fundamentals vs financial stress. *IMF Working Paper, 10*(281).

Bouillot, R., Candelon, B., & Kool, C. (2024). Predicting financial fragmentation using machine learning. *Manuscript in preparation*, Maastricht University.

Candelon, B., Luisi, A., & Roccazzella, F. (2022). Fragmentation in the European Monetary Union: Is it really over? *Journal of International Money and Finance, 122*, Article 102574.

Ciarlone, A., Piselli, P., & Trebeschi, G. (2009). Emerging markets' spreads and global financial conditions. *Journal of International Financial Markets, Institutions and Money, 19*(2), 222–239.

Csontó, B. (2014). Emerging market sovereign bond spreads and shifts in global market sentiment. *Emerging Markets Review, 20*, 58–74.

Csontó, B., & Ivaschenko, I. (2013). Determinants of sovereign bond spreads in emerging markets: Local fundamentals and global factors vs. ever-changing misalignments. *IMF Working Paper, 13*(164).

Edwards, S. (1984). LDC foreign borrowing and default risk: An empirical investigation, 1976–1980. *American Economic Review, 74*(4), 726–734.

Hadzi-Vaskov, M., & Ricci, L. A. (2022). Does gross or net debt matter more for emerging market spreads? *Journal of Money, Credit and Banking, 54*(6), 1777–1802.

International Monetary Fund. (2023). *Global financial stability report: Financial and climate policies for a high-interest-rate era*. Washington, DC.

International Monetary Fund. (2024). *Global financial stability report: Steadying the course: Uncertainty, artificial intelligence, and financial stability*. Washington, DC.

Monetary Policy and Radical Uncertainty

Paul De Grauwe and Yuemei Ji

Introduction

The nature of uncertainty matters a great deal for the conduct of macroeconomic policies in general and monetary policies in particular. This has always been the case. It will probably be even more so in the future when climate change is likely to create existential crises and thus new dimensions of uncertainty for everybody, including policymakers. In this connection the notion of "radical uncertainty" has become popular (see Kay and King (2020)); a kind of uncertainty that will radically affect the nature of policymaking.

To understand how radical uncertainty may affect policies it is necessary to make clear what is meant by "radical uncertainty". Radical uncertainty can be defined in different ways. Here we will consider two definitions. We distinguish between a strong and a soft definition.

The strong definition interprets radical uncertainty in the sense of Frank Knight (Knight (1921)). Uncertainty is radical when we cannot quantify it. In particular, it is impossible to know the frequency distribution of macroeconomic shocks and macroeconomic variables in general. We are in the realm of the "unknown unknowns". Anything can happen. Large shocks can occur, but there is no way of knowing the nature and the timing of these shocks.

There is a softer definition of radical uncertainty. This is a situation where the frequency distributions deviate from Gaussian (normal) distribution. Specifically, they have fat tails ("Black swans", Taleb (2007)) and they may not be one-modal. We will mostly discuss radical uncertainty in this second sense.

Purists may counter that the second definition is not really radical, and this may be true. The advantage of using this soft definition of radical uncertainty, however, is that we can model it, and we can come to some conclusions that deviate from mainstream macroeconomics. But we will occasionally refer to radical uncertainty in the first sense to find out how it interacts with the second one.

How can one model radical uncertainty (in its soft definition)? We will show that radical uncertainty arises because agents fail to understand the underlying model. Thus, one can have an underlying model that is relatively simple, but when we assume that agents do not understand its structure, we obtain complexity and radical uncertainty. There is no need to model complexity to obtain this result.

As will be shown, the intriguing thing is that in a world where agents do not understand the model one creates complexity that is very difficult to understand, thereby validating why agents do not understand the model.

This contrasts with mainstream macroeconomic models that assume Rational Expectations (RE), i.e. that assume that agents understand the underlying model and know the distributions of the shocks. Such a RE-model does not generate radical uncertainty (as defined here in its soft sense). Radical uncertainty can only come from outside the model. Mainstream macroeconomics only recognizes the exogenous shocks as sources of radical uncertainty. Thus, there is no place for endogenous sources of radical uncertainty because agents with rational expectations understand the workings of the underlying model of capitalism. Once the shock has occurred, they can compute with great confidence how it will be transmitted to the economy (for a profound analysis of how modern macroeconomics went wrong, see Stiglitz (2018)).

A simple behavioral macroeconomic model

We use a simple behavioral macroeconomic model. The model's equations are shown in appendix (see also De Grauwe (2012) and De Grauwe and Ji (2019)). It consists of a standard aggregate demand equation, a New-Keynesian supply equation and a

Taylor rule equation. Aggregate demand is a function of expected future demand and the real interest rate. The New-Keynesian supply curve explains the rate of inflation by the expected inflation and the output gap. The Taylor rule describes the behavior of the central bank that manipulates the nominal interest rate to keep inflation close to its target and to stabilize the output gap.

It is assumed that agents do not know the structure of the model in which they operate. The model is too complex to be understood by humans. Therefore, they use simple rules (heuristics) to guide their behavior. These agents are rational, however, in that they are willing to learn from their mistakes. Thus, when they find out that the rule they are using performs less well than alternative rules, they switch rules. This switching rule is a way for agents to learn about the economy.

We will illustrate how in this model radical uncertainty emerges in different forms:

- movements in macroeconomic variables that are not normally distributed and that exhibit fat tails even if the exogenous shocks are normally distributed.
- impulse responses to shocks that are not normally distributed leading to the problem that conditional forecasting ("what if"') cannot be answered properly; it also leads to the problem that policy analyses based on representing the effect of policy actions (e.g. interest rate hikes) by impulse responses cannot be analyzed properly either.

Deviations from normal distributions

We start by presenting basic results of the model. Given the non-linear nature of the switching rules (see equations (4) to (15) in appendix), we have recourse to numerical methods. We simulated the model using numerical values of the coefficients obtained from the literature and imposing i.i.d. shocks in the demand and supply equations and the Taylor rule equation, with zero mean and standard deviations of 0.5. These shocks produce first moments of the output gap and inflation that mimic the first moments found in the empirical data (see Reifschneider and Williams (1999) and Chung, et al. (2012)).

Figure 1 presents the movements of the output gap in the time domain (left panel) and in the frequency domain (right panel). Figure 2 shows the movements of "animal spirits" which express market sentiments of optimism and pessimism generated endogenously in the model (see appendix). It is an index varying between -1 (extreme pessimism) and + 1 (extreme optimism). We observe that the model produces waves of optimism and pessimism (animal spirits) that can lead to a situation where everybody becomes optimist (S_t = 1) or pessimist (S_t = -1).

As can be seen from the left hand side panels, the correlation of these animal spirits and the output gap is high. In the simulations reported in Figure 1 this correlation reaches 0.94. Underlying this correlation is the self-fulfilling nature of expectations. When a wave of optimism is set in motion, this leads to an increase in aggregate demand. This increase in aggregate demand leads to a situation in which those who have made optimistic forecasts are vindicated. This attracts more agents using optimistic forecasts. This leads to a self-fulfilling dynamics in which most agents become optimists. It is a dynamics that leads to a correlation of the same beliefs. The reverse is also true. A wave of pessimistic forecasts can set in motion a self-fulfilling dynamics leading to a downturn in economic activity (output gap). At some point, most of the agents have become pessimists.

The right hand side panels show the frequency distribution of output gap and animal spirits. We find that the output gap is not normally distributed, with excess kurtosis and fat tails. A Jarque-Bera test rejects normality of the distribution of the output gap. The origin of the non-normality of the distribution of the output gap can be found in the distribution of the animal spirits. We find that there is a concentration of observations of animal spirits around 0. This means that much of the time there is no clear-cut optimism or pessimism. We can call these periods of "Great Moderation". The excess kurtosis tells us that there is a high concentration of such periods. There is also, however, a concentration of extreme values at either -1 (extreme pessimism) or +1 (extreme optimism). These extreme values of animal spirits explain the fat tails observed in the distribution of the output gap. The interpretation of this result is as follows. When the market is gripped by a self-fulfilling movement of optimism (or pessimism) this can lead to a situation where everybody becomes optimist (pessimist). This then also leads to an intense boom (bust) in economic activity.

Figure 1: Output gap

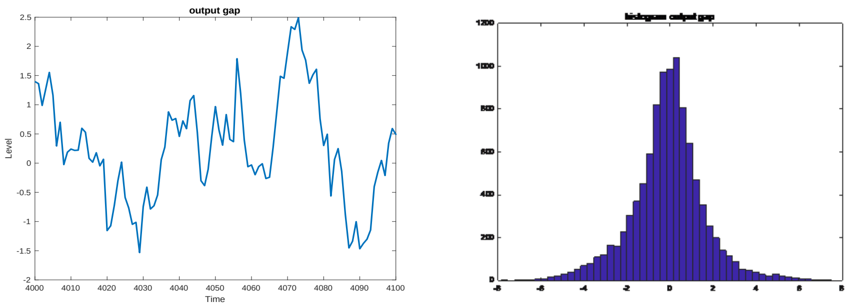
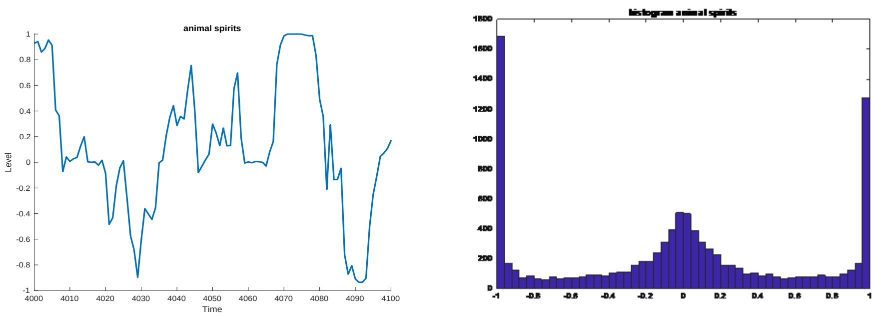


Figure 2: Animal Spirits



Source: De Grauwe and Ji (2019)
Note: the model was simulated over 10,000 periods. The representations in the time domain show a representative sample of 100 periods. The frequency domain shows all periods.

The economy switches from normal periods to extreme movements of booms and busts in an unpredictable way. This is what produces complexity and radical uncertainty, which we defined earlier as deviations from Gaussian distributions. (Note that the exogenous shocks are normally distributed). Two factors explain this complexity. First, there is the ignorance of agents about the underlying model, and second, their attempt to understand by a "trial and error" learning mechanism. The

complexity that is created in this way justifies the assumption that agents have insufficient cognitive abilities to understand the underlying model.
Note also that the switching from Great Moderation to booms-and-bust regimes gives the impression of changes in the structure of the model, although no such structural changes occur. This also adds to the complexity of the dynamics of the model.

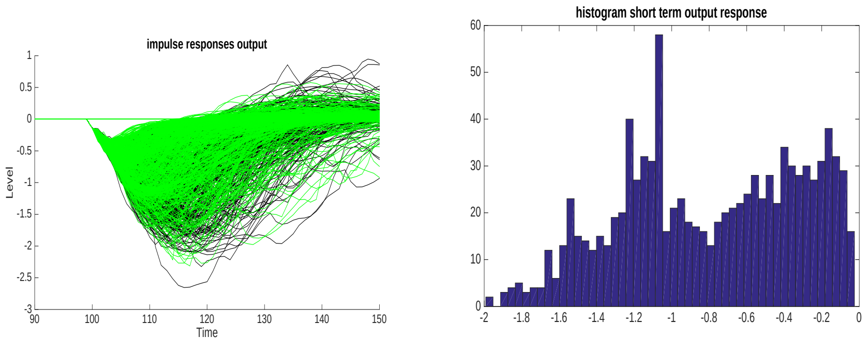
Impulse responses

In contrast to linear rational expectations models the impulse responses depend on the timing of the shock. Put differently, an impulse response computed with one realization of the stochastic shocks in the equations of the model will be different from an impulse response to exactly the same shock but performed using another realization of these stochastic shocks. This is the case even when all parameters of the model are identical.

In order to illustrate this we simulated 1000 impulse responses of the output gap to the same (one standard deviation) negative supply shock occurring at a particular point in time, assuming each time a different realization of stochastic shocks of the model. We show these impulse responses in Figure 3, in the time domain and the frequency domain. We obtain a collection of 1000 impulse responses. Note that the responses in the frequency domain are obtained by collecting these responses 12 periods (3 years) after the supply shock. So, the frequency domain figure is just the intersection of the observations of the time series 12 periods after the supply shock. Several features of these impulse responses stand out.

First, there is sensitivity to initial conditions. We obtain very different impulse responses to the same shock, depending on the initial conditions. The representation in the frequency domain shows that the distribution is not at all Gaussian. It is difficult to infer any structure in this distribution. As a result, it is very difficult to make conditional forecasts about how a negative supply shock will affect the output gap, except that the effect is negative, and that after a sufficiently long period this negative effect will tend to disappear.

Figure 3: Impulse response output gap to a (1 standard deviation) negative supply shock
(b)



Source: De Grauwe and Ji(2024)

The size of the exogenous shocks

In this section we show that increases in the size of the exogenous shocks (demand and supply shocks) have strongly non-linear effects on how output and inflation behave. Put differently, a doubling of the standard deviations of these shocks is not just translated into a doubling of the standard deviations of output and inflation; it changes the dynamics of the volatility in these endogenous variables. To show this, we allowed the standard deviation of the demand and supply shocks to increase from 0.5 to 2, and analyzed how this affected the volatility of output and inflation.

We show the results of applying standard deviations of demand and supply shocks of 2 in Figures 4 and 5. These results should be compared those in Figures 1 and 2 where we assumed a standard deviation of 0.5. We find that in general the volatility of output gap and Inflation (not shown here) multiplies by a factor of 4. The most striking aspect of these results, however, is that it changes the nature of the volatility of these variables. This can clearly be seen from the distribution of the animal spirits. When comparing Figure 5 with Figure 2 we find that there is a much larger concentration of extreme values of animal spirits in the high volatility regime. Most observations on animal spirits are now either extreme optimism or extreme pessimism (+1 or -1). This translates into wild swings in booms and busts in the output gap. Thus, average volatility increases, but this volatility is also of the more extreme variety (fat tails) increasing radical uncertainty.

Figure 4: Output gap

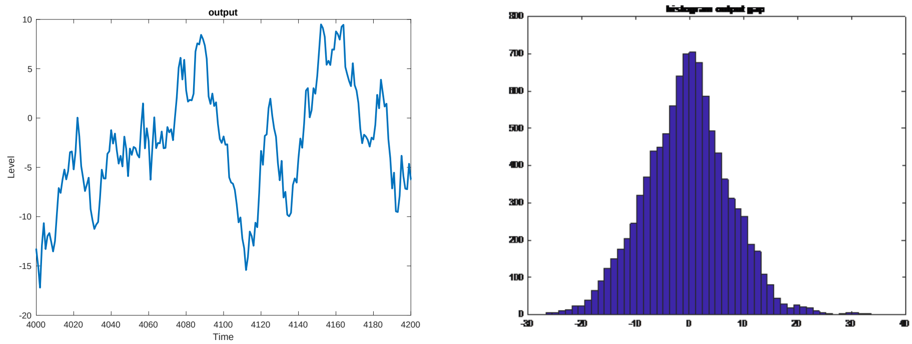
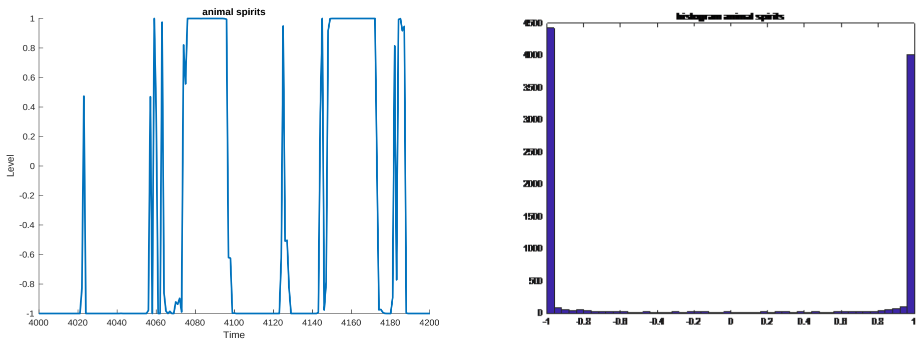
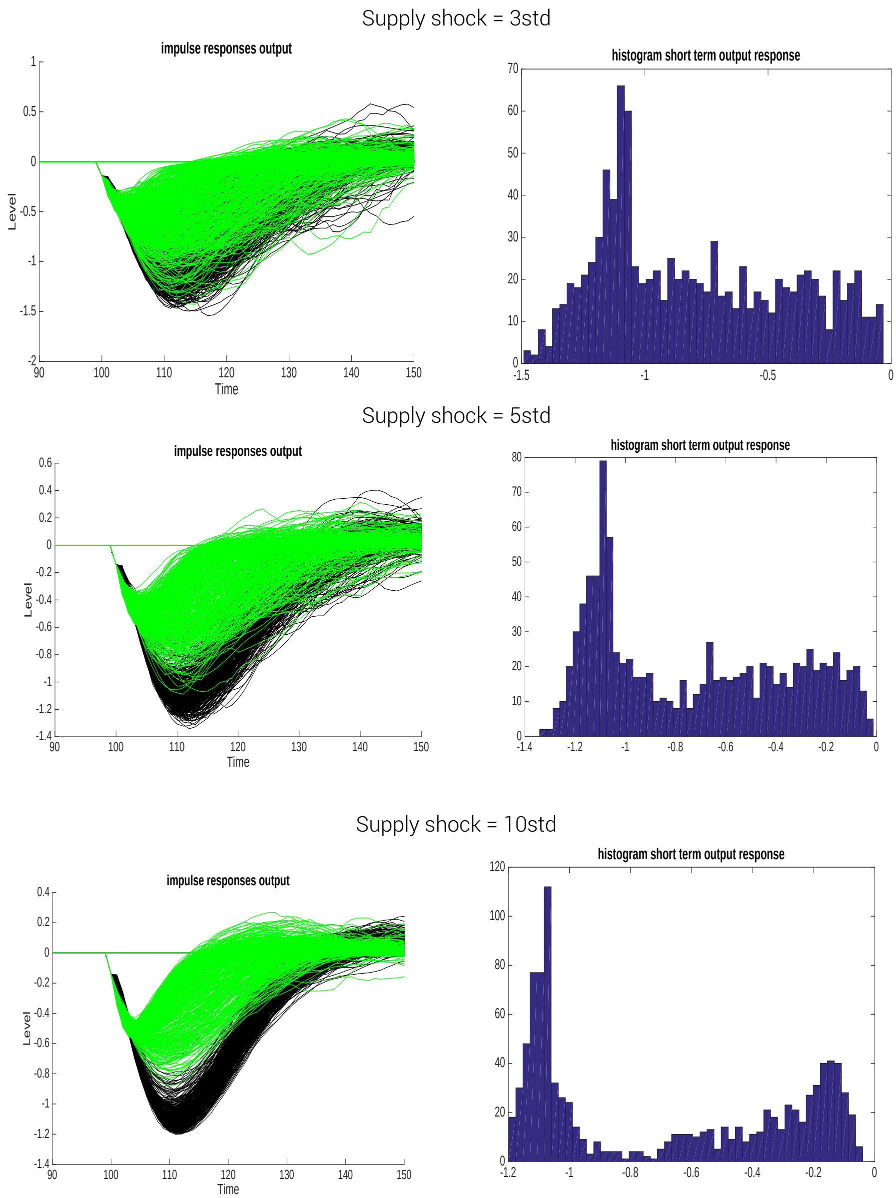


Figure 5: Animal Spirits



We asked a similar question about how the size of the shocks affects the impulse responses. We show this in Figure 6 by increasing the size for the negative supply shock to respectively, 3 std, 5 std and 10 std. A striking result is that by increasing the size of the shock, there appears to be more structure in the distribution of the impulse responses. However, this structure does not show any relation to the normal distribution. What appears is a movement to a bi-modal distribution. This is especially evident when the impulse responses following a very large negative supply shock of 10std. This is the kind of shock experienced during the pandemic of 2020-21, when many countries saw their GDP decline by 10 percent of more. Clearly, this was a shock of the "unknown unknown" type and arises from radical uncertainty in its strong definition. Figure 6 makes clear that such a large shock changes the transmission mechanism of the shock in a fundamental way.

Figure 6: Impulse responses to negative supply shocks of varying sizes



Source: De Grauwe and Ji (2024)

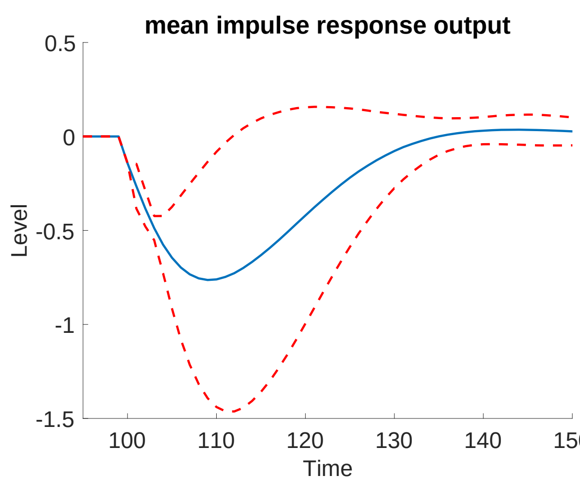
This transmission mechanism now shows a strong bi-modal structure. It appears that there are two types of trajectories taken by the impulse responses when the supply shock is very large. There is a set of "good" trajectories (colored green) which shows the collection of output responses to the negative supply shock that are relatively mild, and a set of "bad" trajectories (colored black) showing a collection of responses that lead to significantly stronger declines in the output gap. The existence of these two trajectories is clearly shown in the frequency distribution of the output gap responses 12 periods after the shock. We observe a concentration of responses around -0.15 and a concentration around -1.1. (These numbers are the multipliers of the supply shocks on output).

One interesting aspect of these results is that these two trajectories depend on the initial conditions. When the latter are favorable, (i.e. initial inflation and inflation expectations are low) we end up in a good trajectory. When in contrast the initial conditions are unfavourable (high initial inflation and inflation expectations) we end up in the bad trajectory characterized by a deep recession that lasts much longer than in the good trajectory.

It is important to understand why we obtain this result. Here is the explanation (for more detail see De Grauwe and Ji(2024)). When a large negative supply shock occurs under unfavorable initial conditions (high initial inflation and low output) the negative supply shock quickly leads to a loss of trust in the central bank's capacity for keeping inflation close to the target and a loss of trust in the central bank's capacity to stabilize output. As a result, few agents will want to use the inflation target as their forecasting rule and many revert to an extrapolative rule. A similar effect holds for the output gap: agents will not have confidence that the output gap will quickly return to its fundamental value. This means that the mean reverting processes in the expectations formations are switched off and only the extrapolating dynamics remains. This creates a destabilizing dynamic that keeps the inflation and output gap high. In contrast when the initial conditions are favorable (low inflation expectations and optimism about the economy) the same negative supply shock does not push credibility and animal spirits against its limits. Mean reverting processes continue to do their work of softening the impact of the supply shock and one ends up in a good trajectory.

Note also that within these two trajectories there is a lot of variation of the particular paths the impulse responses will take. (For more detail on the role of the initial conditions and for an interpretation of these results, see De Grauwe and Ji (2024)). The frequency distributions of impulse responses that we analyzed in the previous sections show strong departures from the Gaussian distribution. As a result, the mean response and the standard deviations of these responses are not informative about the true underlying distribution. We illustrate this problem as follows. We use the impulse responses of output to a 10 std deviation negative supply shock from Figure 6 and compute the mean and the two standard deviations below and above the mean. We show the results in Figure 7.

Figure 7: Mean impulse responses output after large supply shock



Comparing these with Figure 6 the mean and the standard deviations are not only uninformative but even misleading about the true underlying distribution because Figure 7 gives the impression of the existence of a central tendency, the mean, that is representative of the impulse responses. In fact, there are almost no observations close to the mean as the impulse responses are clustered away from the mean. In addition, the representation in Figure 7 gives the wrong impression that, as one moves away from the mean, observations become less likely. In fact, the opposite is true.

This leads to the following problem. A standard assumption made in mainstream RE models is that agents know the distribution of the shocks, typically assumed to be Gaussian. The impulse responses derived from such an analysis typically have a representation as in Figure 7. This only makes sense if the distribution of these responses is Gaussian. If they are not, as is the case in our model, these representations are generally misleading.

The main business of macroeconomists is to produce conditional forecasts i.e. producing mean effects of some shock and a band of uncertainty around the mean. This could be a policy shock, a demand and supply shock, and many others. In a non-Gaussian world these conditional forecasts cannot be trusted. This leads to the idea that when making conditional forecasts one has to think in terms of *scenarios*. There are good scenarios and bad scenarios. In our model the probability of each of these scenarios is 50%. We can, however, make more precise forecasts if we know the initial conditions when the shock occurred (See De Grauwe and Ji (2024)).

In this connection, it is useful to introduce the notion of *ambiguity*. There is strong ambiguity about the effects of shocks because the same shock can lead us into different universes of adjustment. In other words, without the knowledge of initial conditions, the distribution of the impulse responses is ambiguous.

Forecasting and radical uncertainty

Modern macroeconomics has given central stage to forward looking agents. This means that agents are assumed to make decisions based on forecasts of the variables that matter. Consumers, for example, are supposed to base their decision to consume on what they expect their future income to be. Similarly, policymakers in a rational expectations setting are assumed to make decisions based on forecasts of the variables they wish to influence. In this logic central bankers should set the interest rate based on their expectations (forecasts) of future inflation and output gap (or growth rate). See Clarida, Gali, Gertler(2000), Batini, and Haldane (1999), Svensson(1997).

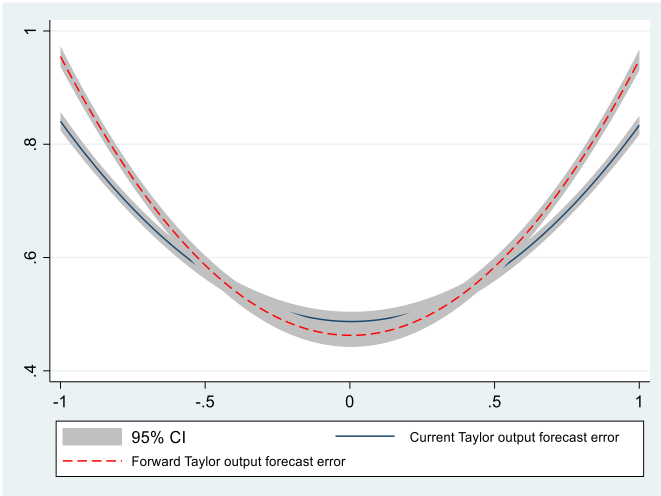
The question that arises here is whether this is a sensible decision rule when the future is radically uncertain. When central banks rely on forecasts to make their decisions, they are likely to often make significant policy mistakes in a world of radical uncertainty. The question then is whether they can improve the quality of their policy decisions by not relying on forecasts of inflation and of the output gap, but rather by relying on currently observed values of these variables.

We analyzed this question in the context of our behavioral macroeconomic model (De Grauwe and Ji (2019)). We used two versions of the Taylor rule equations. The first one uses currently observed values of inflation and output gap. We called this the "current Taylor rule". The second one uses the market forecasts of inflation and output gap. We called this the "forward Taylor rule".

We then simulated the model using i.i.d. shocks in the demand and supply equations and calculated the forecast errors made by agents and by the central bank under the current and forward Taylor rules. We plot the squared forecast errors of output gap (Figure 8) and inflation (Figure 9) against the animal spirits. We find that when animal spirits are close to zero (tranquil times) the forecast errors tend to be the same in the two Taylor rule regimes. As animal spirits increase (in absolute values) the forecast errors increase and more so under the forward-looking Taylor rule.

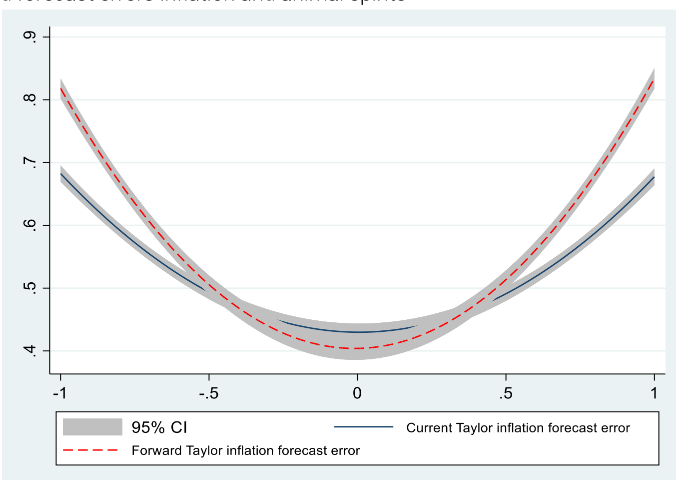
This leads to the following insight. When extreme optimism or pessimism prevails (animal spirits are close to +1 or -1) the economy is in a boom-bust regime with extreme volatility of output and inflation. Given the extreme volatility of these variables when animal spirits are intense, the central bank that uses market expectations will make many policy errors that have to be corrected afterwards. It is then better for the central bank to use currently observed output and inflation to set the interest rate. This leads to lower forecasting errors so that the central bank is likely to make fewer policy errors.

Figure 8: Squared forecast errors output gap and animal spirits



Source: De Grauwe and Ji (2019)

Figure 9: Squared forecast errors inflation and animal spirits



Source: De Grauwe and Ji (2019)

Boom-bust and stabilization

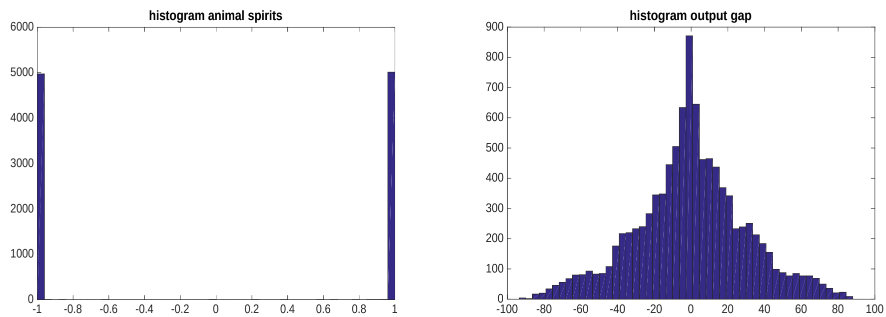
The previous discussion leads to the question of stabilization. We have found out that a world in which agents have cognitive limitations preventing them from having RE is one characterized by frequent boom-bust scenarios that destabilize the economy and that also have the potential of undermining the fabric of society. Can monetary policy do something about this and "stabilize an otherwise unstable" economy? The answer is positive. We simulated our behavioural model assuming normally distributed (and small) shocks. We have seen earlier that under those conditions the model produces regimes of relative tranquillity alternating with occasional bursts of boom and busts. We asked the question of whether the central bank can reduce the occurrence of booms and busts by attaching an increasing weight to output stabilization measured by the parameter c2 in the Taylor rule equation (see equation (3) in appendix). We find that in general this reduces the deviation of the distribution of the variables from normality.

Figure 10 shows how an increase in the stabilization effort reduces the intensity of booms and busts. We show the frequency distribution of animal spirits and the corresponding distribution of the output gap for increasing values of the output parameter (c2) in the Taylor rule equation. We observe several features. First, when c2=0 we have a qualitatively very different result compared with the results obtained when c2>0. This has to do with the fact that when c2=0 we have a chaotic dynamics (see De Grauwe and Ji(2019)). There are then only extreme values of animal spirits and extreme fluctuations of the output. Chaotic dynamics disappears when c2>0.
Second, as c2 increases the frequency with which extreme values of optimism and pessimism occur declines and the concentration around the mean increases. Third, the variability of the output gap declines significantly. This can be seen on the horizontal axis of the distribution of the output gap. With low c2 the output gaps varies between much larger values than when c2 is high. Thus, the intensity of output stabilization has a double effect: it reduces the variability of the output and it reduces the frequency with which extreme booms and busts occur as a result of extreme variation of animal spirits.

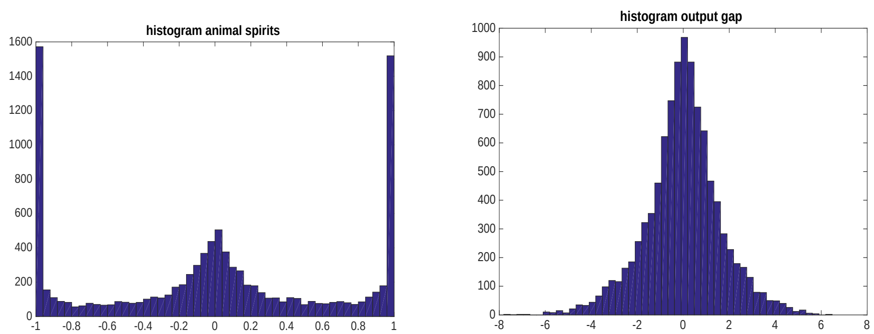
The result of this stabilization effort by the central bank is that fat tails become less fat leading to less intense booms and busts. Thus, there is a role for the central banks to stabilize the business cycle. Just keeping inflation low will not be sufficient. Central banks that pursue strict inflation targeting, without concern for output stabilization, maximize the probability of boom-bust scenarios and the occurrence of fat tails. Paradoxically, as in boom-bust scenarios output becomes very volatile, inflation will also be volatile. It is therefore in the interest of a central bank concerned about price stability to actively stabilize output.

Figure 10: Frequency distribution animal spirits and output gap (c_1=1.5)

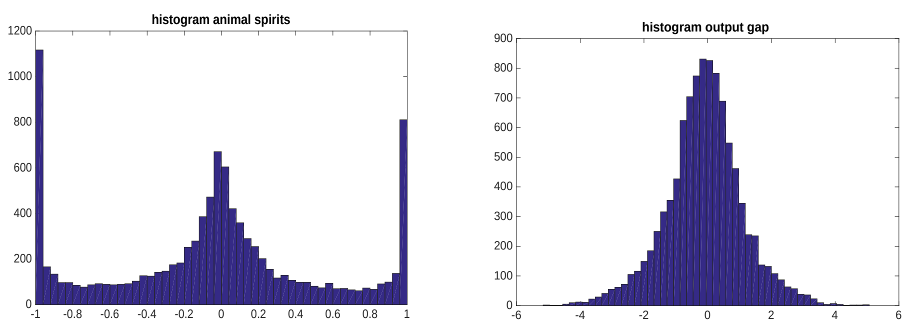
c_2=0



c_2=0.5



c_2=1



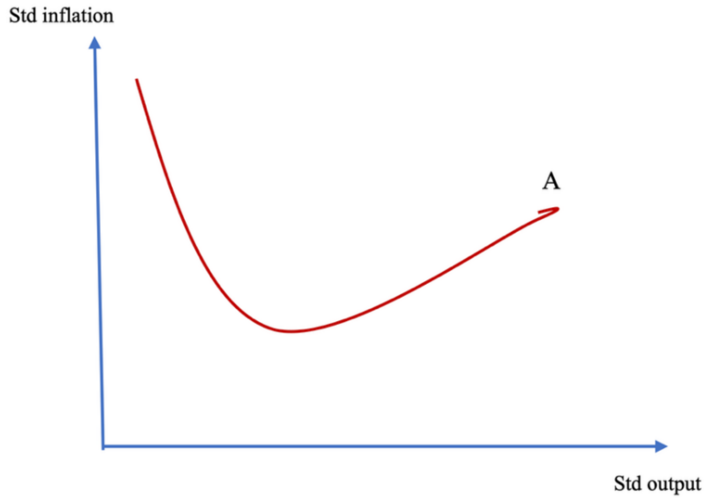
Mainstream DSGE macroeconomics has taken the view that apart from maintaining price stability, the task of the central bank is to reduce the inefficiencies arising from wage and price rigidities. Stabilizing output is motivated by the need to reduce such inefficiencies. The idea that the central bank may be called upon to stabilize an otherwise unstable system is completely absent (Woodford (2003), Smets and Wouters (2003), Gali (2008)).

Here we have shown that there is a need to stabilize a system that is regularly gripped by waves of optimism and pessimism. These waves can lead to violent movements of output and employment. As a result, the need to stabilize runs much deeper and creates greater responsibilities of the monetary authorities than one obtains from mainstream macroeconomics.

Clearly, that does not eliminate trade-offs. In De Grauwe and Ji (2019) we derive the trade-offs between inflation and output stabilization. In contrast to RE models where these trade-offs are negative, (i.e. the pursuit of more output stability always leads to less inflation stability), we find in our behavioral model that this trade-off is non-linear. We show an example in Figure 11. This shows the standard deviation of inflation on the vertical axis and the standard deviation of output on the horizontal axis.

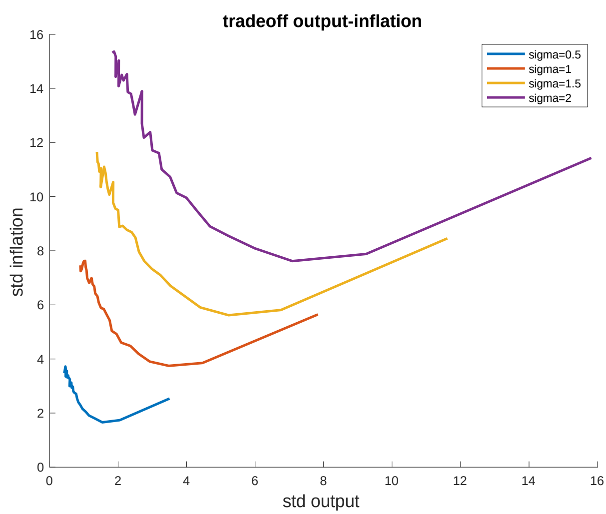
To understand this trade-off start from point A. This is the point where the central bank only pursues price stability, with no effort to stabilize output. When the central bank increases its intensity to stabilize output (by increasing the output coefficient in the Taylor rule) we move downwards on the trade-off curve. This means that by doing more output stabilization, the central bank reduces *both* output and inflation volatility. At some point, however, one hits the minimum point on the trade-off curve. Further attempts to stabilize output will then lead to more inflation volatility. We then reach the standard negative segment on the trade-off. This leads to the conclusion that there is some optimal degree of output stabilization. It also leads to the conclusion that a no-output stabilization strategy is sub-optimal.

Figure 11: Trade-off volatility of inflation and of output



Next, we ask the question of how the size of the demand and supply shocks affects these trade-offs. We simulated our model using progressively increasing standard deviations (sigma) of these shocks from 0.5 to 2. We show the results in Figure 12. Not surprisingly, with higher standard deviations of the shocks, the trade-off curves shift upwards, i.e. in a world of increasing volatility the trade-offs become more unfavourable.

Figure 12: Trade-offs with increasing volatility of demand and supply shocks



More interesting, however, is the result that gains from stabilization measured by the upward sloping segments of the trade-offs increase. Put differently, moving into a regime for more volatility of exogenous shocks makes the need for output stabilization by the central bank more compelling. This has to do with the fact (observed earlier) that in a more volatile environment booms and busts become more intense, increasing the need to "stabilize an otherwise unstable system".

Climate scientists have made it clear that climate change will have dramatic effects on living conditions on our planet. Yet, there is considerable uncertainty about how and when these effects will hit us. There is, in other words, radical uncertainty about how and when the planet will be affected.

In a recent article Annicchiarico, et al. (2024) analyze the macroeconomic implications of climate change using a behavioral macroeconomic model like the one used here, i.e. it is a model where agents face cognitive limitations to understand the complexity of the world, and as a result use simple heuristics to make forecasts. This model produces similar business cycle behavior, and departures from Gaussian distributions as those discussed earlier. These authors find that in such a model it will be more difficult to stabilize the economy and to keep inflation low when climate change occurs, compared to a model where agents are assumed to have Rational Expectations (RE). This is not really surprising. Agents with RE understand the nature of the climate change hitting them and take the necessary precautions in terms of saving and consumption, helping to keep the economy on a steadier path. When agents have cognitive limitations this becomes more difficult to achieve as boom-bust scenarios (fat tails) will undermine the stability of the economy. It is confirmed by the simulations of the tradeoffs presented in this section.

Conclusion

In this paper we have analyzed how radical uncertainty in its various appearances affects the movements of macroeconomic variables. We have argued that in a world of radical uncertainty there will be deviations from normality in the frequency distributions of macroeconomic variables. This then becomes a world of frequency distributions with fat tails. It is also a world in which the transmission of large shocks cannot be forecasted. This led us to question the use of impulse responses. The latter are traditionally used to forecast how shocks are transmitted to the economy. The underlying assumption usually is that the frequency distribution of these impulse responses is Gaussian. In a world of non-normality these impulse responses are misleading.

We analyzed the implications of radical uncertainty for monetary policy. Using a simple behavioral macroeconomic model in which agents have cognitive limitations preventing them from having Rational Expectations, we showed that in such models the central bank bears a much greater responsibility to stabilize an otherwise unstable system than in mainstream models that assume Rational Expectations.
We also showed that when, exogenous shocks hitting the economy get larger it will be more difficult to stabilize the economy and to keep inflation low, not only because the shocks are bigger but also because it will lead to sharper deviations from normality in the distributions of the macroeconomic variables, producing more frequent boom-bust scenarios. The solution to this problem would be that central banks put more emphasis on output stabilization, even if their mandate forces them to deliver on price stability only.

Appendix:
A simple behavioral macroeconomic model

Basic equations

We decided to use the simplest possible behavioral model, (for more complex models see e.g. Delli Gatti, et al. (2005)). The basic model consists of an aggregate demand equation, an aggregate supply equation and a Taylor rule as described by De Grauwe (2012, 2019 and 2020). The aggregate demand and supply equations can be derived from expected utility maximization of consumers and expected profit maximization of firms (Hommes and Lustenhouwer (2019) and De Grauwe and Ji (2019)). In De Grauwe and Ji (2019) we provide a microfoundation.

The aggregate demand equation can be expressed in the following way:

$$y_t = a_1 \widetilde{E}_t y_{t+1} + (1 - a_1) y_{t-1} + a_2 \left(r_t - \widetilde{E}_t \pi_{t+1} \right) + v_t \qquad (1)$$

where y_t is the output gap in period t, r_t is the nominal interest rate, π_t is the rate of inflation and two forward looking components , $\widetilde{E}_t \pi_{t+1}$ and $\widetilde{E}_t y_{t+1}$. The tilde above *E* refers to the fact that expectations are not formed rationally. How exactly these expectations are formed will be specified in the next section.

The aggregate supply equation is represented in (2). This New Keynesian Philips curve includes a forward looking component, $\widetilde{E}_t \pi_{t+1}$, and a lagged inflation variable. Inflation π_t is sensitive to the output gap y_t.

$$\pi_t = b_1 \widetilde{E}_t \pi_{t+1} + (1 - b_1) \pi_{t-1} + b_2 y_t + \eta_t \qquad (2)$$

The Taylor rule describes the central bank's behaviour in setting the interest rate. This behaviour can be described as follows:

$$r_t = (1 - c_3)[c_1(\pi_t - \pi^*) + c_2 y_t] + c_3 r_{t-1} + u_t \qquad (3)$$

where r_t is the interest rate in period t, π_t is the inflation rate, π^* is the target rate of inflation and y_t is the output gap. The error terms in each of the equations (1) to (3) are assumed to be normally distributed with mean zero and a constant standard deviation.

Expectations formation

We analyze how the forecast of output gap $\widetilde{E}_t y_{t+1}$ and inflation $\widetilde{E}_t \pi_{t+1}$ are formed in the model. The rational expectations hypothesis requires agents to understand the complexities of the underlying model and to know the frequency distributions of the shocks that will hit the economy. We take it that agents have cognitive limitations that prevent them from understanding and processing this kind of information. These cognitive limitations have been confirmed by laboratory experiments and survey data (see Branch, 2004; Hommes (2011, 2021)).

Forecasting the output gap

We assume two types of rules agents follow to forecast the output gap. A first rule is called a "fundamentalist" one. Agents use the steady state value of the output gap (which is normalized at 0) to forecast the future output gap. A second forecasting rule is a "naïve" extrapolative one. Following this rule, agents extrapolate the previous observed output gap into forecasting the future. The fundamentalist and extrapolator rules for output gap are specified as follows:

$$\widetilde{E}_t^f y_{t+1} = 0 \qquad (4)$$

$$\widetilde{E}_t^e y_{t+1} = y_{t-1} \qquad (5)$$

This kind of simple heuristic has often been used in the behavioral macroeconomics and finance literature where agents are assumed to use fundamentalist and chartist rules (see Brock and Hommes (1997), Branch and Evans (2006), Brazier et al. (2008)). In De Grauwe and Ji (2019) more complex rules have been experimented with without changing the nature of the results fundamentally.

The market forecast can be obtained as a weighted average of these two forecasts, i.e.

$$\widetilde{E}_t y_{t+1} = \alpha_{f,t} \widetilde{E}_t^f y_{t+1} + \alpha_{e,t} \widetilde{E}_t^e y_{t+1} \qquad (6)$$

$$\alpha_{f,t} + \alpha_{e,t} = 1 \qquad (7)$$

where $\alpha_{f,t}$ and $\alpha_{e,t}$ are the probabilities that agents use the fundamentalist and the naïve rule respectively.

We specify a switching mechanism of how agents adopt specific rule. Using discrete choice theory (see Anderson, de Palma, and Thisse, (1992) and Brock & Hommes

(1997)) to work out the probability of choosing a particular rule (see De Grauwe and Ji (2019) for more detail. We obtain:

$$\alpha_{f,t} = \frac{exp\left(\gamma U_{f,t}\right)}{exp\left(\gamma U_{f,t}\right)+exp(\gamma U_{e,t})} \tag{8}$$

$$\alpha_{e,t} = \frac{exp(\gamma U_{e,t})}{exp\left(\gamma U_{f,t}\right)+exp(\gamma U_{e,t})} \tag{9}$$

where $U_{f,t}$ and $U_{e,t}$[1] are the past forecast performance (utility) of using the fundamentalist and the extrapolative rules. The parameter γ measures the "intensity of choice". It can also be interpreted as expressing a willingness to learn from past performance. When γ = 0 this willingness is zero; it increases with the size of γ.

Forecasting inflation

Agents also forecast inflation using a similar heuristic, with one rule that could be called a fundamentalist rule and the other a naïve extrapolative rule (see Brazier et al.(2008) for a similar setup). In an institutional set-up, the central bank announces an explicit inflation target. The fundamentalist rule will be called an "inflation targeting" rule described in (10), i.e. agents who have confidence in the credibility of the central bank use the announced inflation target to forecast inflation.

$$\widetilde{E}_t^f \pi_{t+1} = \pi^* \tag{10}$$

where the inflation target is π^*. Agents who do not trust the announced inflation target use the naïve rule, which consists in extrapolating inflation from the past into the future. The "naive" rule is defined by

$$\widetilde{E}_t^e \pi_{t+1} = \pi_{t-1} \tag{11}$$

[1]Note $U_{f,t} = -\sum_{k=0}^{\infty} \omega_k \left[y_{t-k-1} - \widetilde{E}_{f,t-k-2} y_{t-k-1}\right]^2$ and $U_{e,t} = -\sum_{k=0}^{\infty} \omega_k \left[y_{t-k-1} - \widetilde{E}_{e,t-k-2} y_{t-k-1}\right]^2$

The market forecast is a weighted average of these two forecasts, i.e.

$$\widetilde{E}_t \pi_{t+1} = \beta_{f,t} \widetilde{E}_t^f \pi_{t+1} + \beta_{e,t} \widetilde{E}_t^e \pi_{t+1} \quad (12)$$

$$\beta_{f,t} + \beta_{e,t} = 1 \quad (13)$$

Where $\beta_{f,t}$ and $\beta_{e,t}$ are the probabilities that agents use the fundamentalist and the extrapolative rules respectively. The same selection mechanism is used as in the case of output forecasting to determine the probabilities of agents trusting the inflation target and those who do not trust it and revert to extrapolation of past inflation. This inflation forecasting heuristics can be interpreted as a procedure of agents to find out how credible the central bank's inflation targeting is. If this is credible, using the announced inflation target will produce good forecasts and as a result, the probability, $\beta_{f,t}$, that agents will rely on the inflation target will be high. If on the other hand the inflation target does not produce good forecasts (compared to a simple extrapolation rule) the probability that agents will use it will be small. Using the switching mechanism similar to the one specified in equations (8) and (9), we can compute the probability of choosing a particular rule.

$$\beta_{f,t} = \frac{exp\left(\gamma U'_{f,t}\right)}{exp\left(\gamma U'_{f,t}\right) + exp\left(\gamma U'_{e,t}\right)} \quad (14)^2$$

$$\beta_{e,t} = \frac{exp\left(\gamma U'_{e,t}\right)}{exp\left(\gamma U'_{f,t}\right) + exp\left(\gamma U'_{e,t}\right)} \quad (15)$$

The probability, $\beta_{f,t}$, that agents will rely on the inflation target to make inflation forecasts can also be interpreted as the fraction of agents who trust the central bank's inflation target.

The forecasts made by extrapolators and fundamentalists play an important role in the model. In order to highlight this role we define an index of market sentiments,

[2] Note $U'_{f,t} = -\sum_{k=0}^{\infty} \omega_k \left[\pi_{t-k-1} - \widetilde{E}_{f,t-k-2}\pi_{t-k-1}\right]^2$ and $U'_{e,t} = -\sum_{k=0}^{\infty} \omega_k \left[\pi_{t-k-1} - \widetilde{E}_{e,t-k-2}\pi_{t-k-1}\right]^2$

which we call "animal spirits", and which reflects how optimistic or pessimistic these forecasts are (see also Franke and Westerhoff (2017)).

The definition of animal spirits is as follows:

$$S_t = \begin{cases} \alpha_{e,t} - \alpha_{f,t} & if\ y_{t-1} > 0 \\ -\alpha_{e,t} + \alpha_{f,t} & if\ y_{t-1} < 0 \end{cases} \tag{16}$$

where S_t is the index of animal spirits. This can change between -1 and +1. There are two possibilities:

When $y_{t-1} > 0$, extrapolators forecast a positive output gap. The fraction of agents who make such a positive forecasts is $\alpha_{e,t}$. Fundamentalists, however, then make a pessimistic forecast since they expect the positive output gap to decline towards the equilibrium value of 0. The fraction of agents who make such a forecast is $\alpha_{f,t}$. We subtract this fraction of pessimistic forecasts from the fraction $\alpha_{e,t}$ who make a positive forecast. When these two fractions are equal to each other (both are then 0.5) market sentiments (animal spirits) are neutral, i.e. optimists and pessimists cancel out and $S_t = 0$. When the fraction of optimists $\alpha_{e,t}$ exceeds the fraction of pessimists $\alpha_{f,t}$, S_t becomes positive.

When $y_{t-1} < 0$, extrapolators forecast a negative output gap. The fraction of agents who make such a negative forecasts is $\alpha_{e,t}$. We give this fraction a negative sign. Fundamentalists, however, then make an optimistic forecast since they expect the negative output gap to increase towards the equilibrium value of 0. The fraction of agents who make such a forecast is $\alpha_{f,t}$. We give this fraction of optimistic forecasts a positive sign. When these two fractions are equal to each other (both are then 0.5) market sentiments (animal spirits) are neutral, i.e. optimists and pessimists cancel out and $S_t = 0$. When the fraction of pessimists $\alpha_{e,t}$ exceeds the fraction of optimists $\alpha_{f,t}$ S_t becomes negative.

The model has non-linear features making it difficult to arrive at analytical solutions. That is why we will use numerical methods to analyze its dynamics. In order to do so, we have to calibrate the model, i.e. to select numerical values for the parameters of the model. In Table 1 the parameters used in the calibration exercise are presented. The values of the parameters are based on what we found in the literature (see Gali (2008) and Blattner and Margaritov (2010)). The model was calibrated in such a way that the time units can be considered to be quarters. The three shocks (demand shocks, supply shocks and interest rate shocks) are independently and identically distributed (i.i.d.) with standard deviations of 0.5%. These shocks produce first moments of the output gap and inflation that mimic the first moments found in the empirical data (see Reifschneider and Williams (1999) and Chung, et al. (2012)).

Table 1: Parameter values of the calibrated model

a1 = 0.5	coefficient of expected output in output equation
a2 = -0.2	interest elasticity of output demand
b1 = 0.5	coefficient of expected inflation in inflation equation
b2 = 0.05	coefficient of output in inflation equation
c1 = 1.5	coefficient of inflation in Taylor equation
c2 = 0.5	coefficient of output in Taylor equation
c3 = 0.8	interest smoothing parameter in Taylor equation
γ = 2	intensity of choice parameter
σ_ε = 0.5	standard deviation shocks output
σ_η = 0.5	standard deviation shocks inflation
σ_u = 0.5	standard deviation shocks Taylor
ρ = 0.5	measures the speed of declining weights in mean squares errors (memory parameter)

Note: the parameter values used here are similar to the ones obtained in Kukacka and Sacht (2022).

References

Anderson, S., de Palma, A., & Thisse, J.-F. (1992). *Discrete choice theory of product differentiation*. MIT Press.

Anniccharico, B., Di Dio, F., & Dilusio, F. (2024). Climate actions, market beliefs, and monetary policy. *Journal of Economic Behaviour and Organization, 219*, 176–208.

Batini, N., & Haldane, A. (1999). Forward-looking rules for monetary policy. In *Monetary policy rules* (pp. 157–202). University of Chicago Press.

Blattner, T., & Margaritov, N. (2010). Towards a robust policy rule for the euro area. *ECB Working Paper Series*, No. 1210.

Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal, 114*(497), 592–621.

Branch, W., & Evans, G. (2006). Intrinsic heterogeneity in expectation formation. *Journal of Economic Theory, 127*, 264–295.

Brazier, A., Harrison, R., King, M., & Yates, T. (2008). The danger of inflation expectations on macroeconomic stability: Heuristic switching in an overlapping generations monetary model. *International Journal of Central Banking, 32*, 2428–2452.

Brock, W., & Hommes, C. (1997). A rational route to randomness. *Econometrica, 65*, 1059–1095.

Chung, H., Laforte, J. P., Reifschneider, D., & Williams, J. C. (2012). Have we underestimated the likelihood and severity of zero lower bound events? *Journal of Money, Credit and Banking, 44*(s1), 47–82.

Clarida, R., Gali, J., & Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics, 115*(1), 147–180.

De Grauwe, P. (2012). *Lectures on behavioral macroeconomics*. Princeton University Press.

De Grauwe, P., & Ji, Y. (2019). *Behavioural macroeconomics: Theory and policy*. Oxford University Press.

De Grauwe, P., & Ji, Y. (2021). On the use of current or forward-looking data in monetary policy: A behavioural macroeconomic approach. *Oxford Economic Papers*.

De Grauwe, P., & Ji, Y. (2024). Trust and monetary policy. *Journal of Forecasting*.

Delli Gatti, D., Di Guilmi, C., Gaffeo, E., Giuloni, G., Gallegati, M., & Palestrini, A. (2005). A new approach to business fluctuations: Heterogeneous interacting agents, scaling laws, and financial fragility. *Journal of Economic Behavior and Organization, 56*, 489–512.

Franke, R., & Westerhoff, F. (2017). Taking stock: A rigorous modelling of animal spirits in macroeconomics. *Journal of Economic Surveys, 31*(5), 1152–1182.

Galí, J. (2008). *Monetary policy, inflation, and the business cycle*. Princeton University Press.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control, 35*(1), 1–24.

Hommes, C. H., & Lustenhouwer, F. (2019). Inflation targeting and liquidity traps under endogenous credibility. *Journal of Monetary Economics, 107*, 37–48.

Hommes, C. H. (2021). Behavioral & experimental macroeconomics and policy analysis: A complex systems approach. *Journal of Economic Literature, 59*(1), 149–219.

Kay, J., & King, M. (2020). *Radical uncertainty: Decision-making beyond the numbers*. Norton.

Knight, F. H. (1921). *Risk, uncertainty, and profit*. Houghton Mifflin Company.

Kukacka, J., Jang, T. S., & Sacht, S. (2022). Estimation of heuristic switching in behavioral macroeconomic models. *Journal of Economic Dynamics and Control*.

Pfajfar, D., & Zakelj, B. (2014). Experimental evidence on inflation expectation formation. *Journal of Economic Dynamics and Control, 44*, 147–168.

Reifschneider, D., & Williams, J. (1999). Three lessons for monetary policy in a low inflation era. Board of Governors of the Federal Reserve.

Smets, F., & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model. *Journal of the European Economic Association, 1*, 1123–1175.

Stiglitz, J. (2018). Where modern macroeconomics went wrong. *Oxford Review of Economic Policy, 34*(1–2), 1–20.

Svensson, L. E. (1997). Inflation forecast targeting: Implementing and monitoring inflation targets. *European Economic Review, 41*(6), 1111–1146.

Taleb, N. (2007). *The black swan*. Random House.

Woodford, M. (2003). *Interest and prices*. Princeton University Press.

Inflation in the context of imported price shocks and markups

Huub Meijers, Joan Muysken and Tom van Veen

Preface

In 1991, Clemens started to work at the department of Economics at the then Faculty of Economics of the Rijksuniversiteit Limburg. Clemens was educated as an econometrician with a keen eye for economics. He was trained at the Erasmus University in Rotterdam and had worked for the Dutch Central Bank in Amsterdam. The move from the Randstad to Maastricht was more than just a geographical move. It was also an "economics-cultural" move or even a shock. Although Clemens had the perfect background to match the need of the department, there was one challenge, both for Clemens and for the department. The challenge was to also match the Rotterdam and the Groningen views on economics. In Rotterdam, Clemens was trained in the free markets, Friedmanian monetarist tradition, implying a strong focus on equilibrium economics and a strong separation between fiscal and monetary policy. In Maastricht, Clemens came to work with macroeconomists that were educated in the (Post-)Keynesian tradition and disequilibrium economics. In this view, free markets are not without failures. Governments have a strong role in "repairing" the many market failures that occur in a free-market economy and both fiscal and monetary tools should be actively employed. These differences in analyses about how to solve problems like unemployment, financial crises and inflation were heavily debated in animated discussions during the daily lunches in the mensa. Both Clemens and his colleagues were quite strong in their opinions, but over the years we have seen a growing consensus between their views, and it occurred that this economics-cultural "clash" was a blessing for the department because it resulted in a productive, fruitful and long-standing cooperation.

Clemens, over the years we have seen you shifting towards the more post-Keynesian view (and yes you managed to get your colleagues to shift as well), but you are not there yet! Let us say that we are both halfway. To honor your numerous merits and achievements for the department and to once again stress the relevance of post-Keynesian analyses, we have chosen to discuss a topic that is close to your economic heart: inflation. This is of course not a monetary phenomenon as we show in this chapter. We analyze the distributional effects of inflation that occurred after the recent rise in import prices. And as it should be, in a proper stock-flow consistent post-Keynesian model.

Introduction

The rise in inflation after the COVID19 period and the start of the war in Ukraine, provoked a debate about profit inflation. It has been argued that after the rise in the price of imports, firms have increased markups and hence increased the profit share. The primary goal of this chapter is to present a model to illustrate how an increase in the price of imports can trigger domestic inflation and to explore whether this inflation is linked to rising markups or rising profits.

We develop a stock-flow consistent model in the tradition of post-Keynesian economics to analyze the relation between import prices, the markup, the profit share and inflation. The model is a three-sector model with households, firms and the foreign sector. There is no financial sector and no government sector. The model analyses an open economy of firms that produce and sell goods, invest, import and export and households that consume and save. Firms set prices as a markup on marginal costs. Households are categorized into two groups: workers and capitalists, with the latter also referred to as rentiers. Building on the seminal work of Kalecki (1971) workers spend their entire income, while rentiers save a portion of their income. Excess savings from the foreign sector or alternatively, imbalances on the current account are absorbed by the domestic firms by foreign direct investments, i.e. via selling bonds to the foreign sector or buying foreign bonds. The aim of the model is to analyze the effects of an increase in import prices and/or the markup of firms on the profit share and on inflation. We first assume the markup to be constant and thereafter we assume that the markup depends on the utilization rate.

Our main findings are first, the observation of an increase in the profit share following an increase in import prices does not necessarily imply an increase in the markup of firms. Even with a constant markup, the profit share can increase. Second, our results show that inflation that is caused by an increase in import prices leads to an increase in the profit share and thus to a change in the distribution of income. The steady state value of the (new) profit share does not depend on the constancy or variability of the markup. Third, by using a variable markup that is positively related to the rate of capacity utilization, we show that import inflation reduces the markup and consequently reduces the profit share of firms. However, in the medium to long run, the profit share converges to the same level as in the case of a constant markup. For an extended period, domestic inflation remains below the rate of inflation observed with a constant markup. Fourth, in our model we find that an increase in the markup (ceteris paribus) has similar effects on the economy as an increase in import prices.

The chapter is set up as follows. In section 2 we survey the empirical literature on profit inflation. We focus on the post-COVID period. From this survey, we conclude that there is confusion about whether an increase in the profit share necessarily coincides with an increase in the markup of firms. Therefore, in section 3, we elaborate on the relation between the profit share and the markup and show that after an increase in import prices, the profit share can increase while the mark up remains constant. To delve further in the role of the markup, we develop a model in which the markup varies positively with the utilization rate, in section 4. We simulate our model to analyze the effects of import price shocks in section 5. Section 6 concludes.

Survey of the literature on profit inflation

The initial increase in inflation after the COVID episode, which was widely ascribed to supply bottlenecks, was accompanied by an increase in the profit share of firms in many countries (see Gerinovics and Metelli (2024) for an overview). This induced many economists, both in policy institutions and in academics, to conclude that firms were using inflation to increase their markups - that is inflation was profit driven sellers inflation (Arce et al., (2023); Capolongo et al., (2023); Glover et al., (2023b); Hansen et al. (2023); Storm, (2022); Weber and Wasner, (2023) – and both Lagarde as president of the ECB and Lane as chief economist of the ECB, quoted in Storm, (2023)). However, as was pointed out by both Colonna et al. (2023) and Lavoie (2023), the increased profit share does not necessary imply that the markup did rise. Consequently, one should consider the development of markups over time to investigate whether inflation was profit-driven.

Various authors have investigated the development of markups over time. For instance Storm (2023) observes an increase in average markups for the US (with a peak in the last observation 2022-4) and concludes: *"Higher markups have, therefore, added to the recent surge in U.S. inflation and (large) American firms have taken advantage of the situation, exactly as has been argued by Lagarde and Lane for the Eurozone."* [1] – see also Glover et al. (2023a), Konczal and Lusiani (2022), Nikiforos and Grothe (2024) and Setterfield (2022).

An interesting observation was made by Gerinovics and Metelli (2023) who argue that *"in aggregate, the increase in US markup was temporary and, since mid-2022, it has been receding towards pre-pandemic levels. … most of this increase is attributed to the dynamics of high markup firms, particularly those firms that have a markup above the 75th percentile in the distribution of all firms … A sectoral decomposition shows that the increase was driven by a few sub-sectors, such as mining and energy, transportation, and healthcare [..], namely those more strongly affected by both the pandemic and the Russian war on Ukraine."*

[1] He adds: "This is not a surprise. Some CEOs have admitted on shareholder calls and in surveys that they have been taking advantage of inflation to raise profit margins by increasing prices beyond what is needed to offset any increase in their input costs (Ferguson and Storm 2023)."

This finding is in line with Acharya et al. (2024) for the Eurozone. They argue that the pandemic lockdowns forced some specific firms *"to increase their prices to maintain their profits (Kalemli-Ozcan et al. (2022))."* These specific price increases were experienced by households as general price increases and therefore households' inflation expectations increased. This induced firms with high pricing power to increase their prices disproportionally, even after supply-chain disruptions eventually eased.

The findings of Acharya et al. are supported by a lot of empirical evidence. A similar observation is made in Cucignatto et al. (2023).[2] However, in both cases the data run till 2022. It may be very well possible that similar to the findings for the US reported above, the situation has normalized again. But, as Gallo (2024) observes:[3]

"Does that mean that all's well that ends well? Not quite. The rate of change of prices is going down, but their level is permanently higher. In the process, wage earners have lost purchasing power, while profit earners have been able to defend and sometimes increase their profit margins. In other terms, the post-COVID recovery has constituted yet another episode of perverse redistribution from wages to profits. Monetary policy can do little to reverse that, and most likely higher rates will contribute to making the process starker. Incomes policies — long forgotten in policymaking and policy debates — could be a valuable tool to foster real wage growth, allowing for an adequate increase in the wage share". (Gallo, 2024)

---

[2] An exception is De Keyser et al. (2023) who do not find evidence for profit inflation in Belgium. Also Hahn (2023) is cautious about the Euro Area – inflation is mainly due to increases in intermediary costs.
[3] We do not support the alternative view expressed by Gerinovics and Metelli (2024):
In the context of the debate on profit-led inflation and greedflation, these results should be interpreted cautiously for two reasons. First, markups may have increased for other reasons than firm greediness, such as firms recouping pandemic losses or because of the uncertainty about future cost increases and menu costs. Second, the term greedflation seems semantically misleading in a market system that builds upon profit-maximising behaviour as a normal feature: if the markup increases, as consumers are willing to pay more (for the same product), then reaping this willingness resembles profit maximising behaviour – rather than greediness – possibly in response to (temporary) changes in consumer preferences.

Pressman (2023) provides a different explanation for the surge in inflation in recent years: the decrease in productivity. He argues that productivity decreased strongly in both the US and Canada in the last two years, while wages did not decrease that strong. With a constant markup on average costs, this would imply inflation. This decrease in productivity might be temporary because of labor hoarding to compensate for temporary labor shortages and COVID aid supporting inefficient firms. It would become more permanent if "*Covid changes the mind set of workers. People are no longer willing to work as hard as possible to succeed. They will accept lower pay to enjoy life more.*"[4]

A final interesting contribution is provided by Bernanke and Blanchard (2023, 2024). In their 2023 paper they focus on the impact of a tight labor market on US inflation. To motivate the Phillips curve they introduce a traditional wage-price setting two equation model, which is used in many textbooks. They add three important features to the model: explicit equations for short- and long-term inflation expectations and they allow for various kinds of supply shocks in price setting. They find:

"most of the inflation surge that began in 2021 was the result of shocks to prices given wages. These shocks included sharp increases in commodity prices, reflecting strong aggregate demand, and sectoral price spikes, resulting from changes in the level and sectoral composition of demand together with constraints on sectoral supply".

But then they continue *"although tight labor markets have thus far not been the primary driver of inflation, we find that the effects of overheated labor markets on nominal wage growth and inflation are more persistent than the effects of product-market shocks. Controlling inflation will thus ultimately require achieving a better balance between labor demand and labor supply."* This latter observation has been strongly criticized by Storm (2024) who argues convincingly that this observation is not supported by their estimation results. This also holds for the results of estimations by 10 Central Banks in leading world economies, summarized in Bernanke and

[4] Pressman (2023, p. 3) It is not clear why this would imply "continued inflationary pressures.".

Blanchard (2024). Hence the extent to which Bernanke and Blanchard argue that an increase in unemployment should be accepted to control inflation is far beyond the results of their model.

Profit inflation in the Netherlands

For the Netherlands, the RABO bank (Erken et. al, (2023)), the CPB Netherlands Bureau for Economic Policy Analysis (Soederhuizen et al, (CPB, 2023)) and the Dutch Central Bank (Hebbink and Ozturk, DNB (2023)) have investigated whether profit inflation has indeed occurred in the Netherlands.

Erken et. al (2023) use a macroeconomic approach and analyze the change in the net operating surplus of companies in the Netherlands between in 2021Q4 and 2022Q4. This surplus has increased by 10.9 billion Euro. They show that sellers prices have increased much stronger than the prices for intermediary goods. Second, the increase in the net surplus was twice as high as the increase in wage costs.

The CPB and DNB use an alternative approach to investigate the existence of profit inflation. They analyze the change in the price index. DNB analyses the GDP index – in this index, import prices are not explicitly included. The DNB analysis shows first that in comparison to earlier years, in 2021 and in 2022, the contribution of net profits to increases in the GDP deflator has increased strongly. Second, the contribution has decreased again in 2023. Third, like in other countries, four sectors have been mainly responsible for the increased contribution: energy, real estate, government and the financial sector (see also Gerinovics and Metelli, (2023)). However, for financial services and in building and reconstruction, profits have decreased in that period. DNB also investigates the development of the macroeconomic profit margin, defined as the ratio of the GDP deflator and the wage costs and concludes that this margin has increased in 2022. Hence the price of GDP has increased stronger than the wage costs, implying an increase in profits. The CPB reaches similar conclusions as DNB in their analysis of the GDP deflator. In 2022 profits were the main driver behind the increase in this deflator.

The CPB also analyses changes in the consumer price index (CPI), which includes taxes. Decomposition of the change in the CPI allows to assess the contribution of wages, profits, net taxes and import prices to changes in the CPI. The analysis of the development of the CP shows that in 2022, increasing import prices were the main drivers behind inflation, followed by profits.

Based on the results of this research, it is difficult to conclude that profit inflation has occurred in the Netherlands. As all studies mentioned above emphasize, a proper measurement of profits is difficult and subject to measurement errors. Moreover, both CPB and DNB show that the contribution of profits to inflation decreases if the sectors energy, raw materials and water are not taken into account – not all firms have managed to increase profits.[5] Finally, when prices are set as a markup on average costs, profits rise when the price of inputs increases, even in case the markup remains constant, as observed above.

However, it is important point to note that in the Netherlands the wage share in GDP has decreased and the profit share has increased over the last decades. To what extent this has been due to a decrease in productivity, profit inflation, a weakening bargaining position of the unions or increasing market power of firms (Eeckhout, (2021)) is an empirical matter and beyond the scope of the chapter.

The relation between profit share and markup

As mentioned in the previous section, both Colonna et al. (2023) and Lavoie (2023) have shown that an increased profit share does not necessarily imply that the markup has increased: the profit share can increase while the markup does not change. We reproduce their analysis below to show how an increase in the price of intermediary inputs such as imports, affects the profit share in case of a constant markup. For this, we must first elaborate on the difference between the prices of output and GDP.

[5] For the Euro area, Hahn (2023) reaches similar conclusions in her analysis of the GDP price index.

Output, imports and value added

The distinction between output and value added is imperative when investigating the impact of price changes in intermediate goods on the profit share. Following e.g. Colonna et al. (2023), real output Q is assumed to be produced according to a Leontief production function with labor L, capital K and imports M as inputs. Capacity output Q^c is defined by:

$$Q^c = \frac{1}{\kappa}K \tag{1}$$

and assuming efficient use of factors of production demand for labor and imports is given by

$$L = \frac{1}{\lambda}Q;\ M = \ \mu Q \tag{2}$$

respectively. In these equations λ is labor-productivity (Q/L), μ the intensity of material goods used in gross production, and κ the capital-output ratio. We explain below in equation (20) how output *Q* is determined by final demand. The ratio between output and capacity output then determines the utilization rate u:

$$u = \frac{Q}{Q^c} \tag{3}$$

Real value-added *Y* is derived from output *Q* by subtracting imports M:[6]

$$Y = Q - M = \frac{\pi + wL}{p_y} \tag{4}$$

Here p_y is the value-added price and w is the nominal wage rate.[7] The right-hand side of equation (2) follows from the definition of nominal gross profits π as output minus total costs:

$$\pi = p_q Q - p_\mu M - wL = p_q Q - TC \tag{5}$$

where $TC(= wL + p_\mu M)$ stands for total costs. Here p_q is the output price and p_μ is the domestic price of imports – the foreign price level times the nominal exchange

---

[6] In the words of Arrow (1974) "Metaphorically, we can imagine capital and labor cooperating to produce an intermediate good, real value added (V), which in turn cooperates with materials to produce the final product." (p. 5). As value added is not measurable directly, statistical offices apply the double-deflation method to determine the price index of value added. Implicitly this puts constraints on the shape of the production function, see e.g Arrow (1974); Sims (1969). Though directly related to the theme discussed here, we consider measurement issues outside the scope of this chapter.

[7] Since GDP can be measured by value added, p_y is also the GDP price.

rate (the latter is assumed to be constant). Since production is proportional to inputs –cf. equation (2)– marginal costs are independent of output and given by $MC = w/\lambda + p_\mu \mu$.

We assume that the price of output is determined by a constant markup m on marginal costs:

$$p_q = (1+m)MC = (1+m)(w/\lambda + p_\mu \mu) \tag{6}$$

From (4) it follows that if the wage rate, labour productivity and the import share in output are constant, the price of output is positively related to the (domestic)price of imports p_μ, given a fixed markup.

The value-added deflator p_y follows from dividing nominal value added by real value added. When we measure nominal value added by profits plus the wage sum, we find combining equations (4) and (5):

$$p_y = \frac{p_q Q - wL - p_\mu M + wL}{Q - M} = \frac{p_q - p_\mu \mu}{1 - \mu} \tag{7}$$

From (7) it follows that the price of output is a weighted average of the value-added deflator and import prices:

$$p_q = (1-\mu)p_y + \mu p_\mu \tag{8}$$

Using (6) we can write the value-added deflator as function of wage rate and import prices:

$$p_y = \frac{(1+m)w/\lambda + m p_\mu \mu}{1 - \mu} \tag{9}$$

It follows directly from comparing (8) and (9) that the value-added deflator and the price index of output are equal in absence of imports, i.e. if μ is zero, or if the value-added price and the price of imports are equal ($p_y = p_\mu$).[8] However, if this is not the case, the value-added deflator and the price index of output are different – this point

[8] See also Sims (1969), p. 470, for a similar result.

is not observed by Hein (2012) and Lavoie (2023). The importance of this point will be elaborated in section 4.3 below.

The impact of markup and import prices on the profit share

Gross profit share in value added is defined as gross profits divided by value added:

$$h = \frac{\pi}{p_y Y} = \frac{p_q Q - TC}{p_q Q - TC + wL} = \frac{(1+m)MC - MC}{(1+m)MC - MC + wL/Q} \quad (10)$$

We note that average costs TC/Q are equal to marginal costs since $TC/Q = MC = w/\lambda + p_\mu \mu$.

The wage share in total costs is:

$$\frac{wL}{TC} \equiv a = \frac{w/\lambda}{w/\lambda + p_\mu \mu} \quad (11)$$

This yields the following expression of the profit share in value-added:

$$h = \frac{m.MC}{m.MC + w/\lambda} = \frac{m}{m + \frac{w/\lambda}{w/\lambda + p_\mu \mu}} = \frac{m}{m + a} \quad (12)$$

From equation (12) it follows that the profit share is not only dependent on the markup and the wage rate, but also on the price of imports. We find:

$$\frac{\partial h}{\partial m} > 0, \qquad \frac{\partial h}{\partial p_\mu} > 0, \qquad \frac{\partial h}{\partial w} < 0 \quad (13)$$

Both an increase in the markup and an increase in the domestic price of imports increase the profit share in GDP- see also Colonna et al. (2023) and Lavoie (2023). An increase of the wage rate has a negative impact on the profit share. Note that equations (12) and (13) imply that at a constant mark-up, an increase in import prices increases the profit share. This finding has created a lot of confusion in the discussion on profit inflation, as we observed in section 2.

The model with a constant markup

In the previous section we show that at a constant markup, an increase in import prices increases the profit share. This implies that an increased profit share does not automatically imply an increase in markup behavior of firms. The question is to what extent this finding depends on the assumption of a constant markup. Alternatively, what is the impact of import prices on the profit share in case of a variable markup? To address this question, we introduce a variable markup and develop a more elaborate model of the economy. We use a stock-flow consistent model to analyze the relation between import prices, the markup and inflation – see for a recent example of this type of modelling Meijers and Muysken (2024). Since we want to be able to derive an analytical solution, we chose for a very simple model inspired by Hein (2012, Ch. 3).The model analyses an open economy of firms that produce and sell goods, invest, import and export and households that consume and save. Firms set prices as a markup on marginal costs. In this model we initially assume that the markup is fixed, a feature that is relaxed in the subsequent section below. Households are categorized into two groups: workers and capitalists, with the latter also referred to as rentiers. Building on the seminal work of Kalecki (1971) workers spend their entire income, while rentiers save a portion of their income. Savings are determined by the savings of rentiers and the retained profits of firms. We add a foreign sector in the most simple way. Firms import and export goods and receive interest income from foreign direct investments. A current account surplus implies a deficit of the foreign sector which is entirely financed by foreign direct investments of domestic firms. Thus, in the model, there are no flows of reserves between the domestic country and foreign countries.

We first assume a constant markup to be able to solve the model analytically. We then find a steady state solution where the utilization rate varies with the mark-up rate and import prices. This enables us to understand better the interaction between the profit share and the markup, when the markup varies pro-cyclically – we elaborate on this in section 5.

Households, firms and the foreign sector in an open economy

We distinguish between three sectors: the household sector, the firm sector and the foreign sector. Households consist of workers and rentiers. Workers consume their entire income; rentiers consume and save part of their income. These savings are invested in bonds B_F issued by firms – this constitutes household wealth W_H. Firms issue bonds to households and use their retained profits for investment in capital $p_q.K$ and to buy foreign bonds B_A , i.e. foreign investments. Wealth of firms then is given by $W_F = p_q.K + B_A – B_F$. We assume a real trade balance surplus as a constant fraction of real output. Firms receive interest income, and any current account surplus is financed by foreign investment. The accumulated foreign debt B_A then constitutes (negative) foreign wealth W_A. The balance sheets of the three sectors are summarized in Table 1.

Table 1 Balance sheet

	Households	Firms	Foreign	
Bonds Firms	B_F	$-B_F$		0
Bonds Foreign		B_A	$-B_A$	0
Capital stock		$p_q.K$		$p_q.K$
Wealth	W_H	W_F	W_A	W_{TOT}

Households receive wage income and interest income. The latter stems from buying bonds from firms, B_F. Total interest income is equal to the nominal rate of interest i times the lagged nominal value of bonds, $i.B_{F,-1}$.[9] The rate of interest is exogenous. Interest income is transferred to the capitalists (rentiers).
Rentiers consume and save a portion of their income. Total savings by households S_H are thus the savings by rentiers and are equal to the saving rate s_h times the income of rentiers:

$$S_H = s_h \, i \, B_{F,-1} \tag{14}$$

Firms produce and sell output and import and export goods. They pay wages and interest to households. Next to that they earn profits and, assuming a trade balance surplus, receive interest income from investment in foreign bonds. Firms' savings S_F are equal to retained firm profits and are equal to gross profits plus interest received on foreign direct investments minus interest payments to rentiers:

$$S_F = \pi + i \, (B_{A,-1} - B_{F,-1}) \tag{15}$$

[9] We use the past value of bonds (and later capital), since stocks are measured end of period.

Firms' savings are invested in foreign bonds B_A and in firm capital $p_q.K$. We ignore depreciation, hence $I = \Delta K$. Investments I, modelled by the growth rate of capital g_K, depend on a fixed parameter (α), the rate of capacity utilisation (u) and the real interest rate $(i - \hat{p}_q)/(1 + \hat{p}_q)$ such that:

$$g_K = \frac{I}{K_{-1}} = \alpha + \beta u_{-1} - \theta \left(\frac{i - \hat{p}_q}{1 + \hat{p}_q} \right) \quad (\alpha, \beta, \tau, \theta > 0) \tag{16}$$

where a hat refers to growth rates. Investment increases as the rate of capacity utilisation increases (if demand increases such that more of the existing capital stock is used, firms start to invest more). Firms invest less if the interest rate increases, that is, if the cost of lending increases.

Investment is financed by bonds issued to households and by retained profits (firm savings) – the remaining retained profits are used for direct investments abroad ΔB_A. We assume flexible bond markets to ensure that firms issue sufficient bonds to absorb household savings and firms buy sufficient foreign bonds to absorb domestic excess savings ($S_A = S_F + S_H - I$). Hence:

$$\Delta B_F = S_H \text{ and } -\Delta B_A = S_A \tag{17}$$

The foreign sector finances its current account deficit by providing bonds B_A to firms. The savings of the foreign sector then are equal to the domestic current account balance and given by:

$$S_A = -\left(p_q E - p_\mu M \right) - iB_{A,-1} \tag{18}$$

Both real imports *M* and real exports *E* are assumed to be a fixed proportion of real output Hence:

$$M = \mu . Q \text{ and } X = \epsilon . Q \tag{19}$$

Note that $S_A$can be negative (in case of a domestic current account surplus) or positive (in case of a domestic current account deficit). The assumption that current account balances are financed purely by bonds is made to avoid the introduction of a banking sector.[10] The implied accumulation of foreign debt held by the domestic economy and

[10] For a more elaborate model see Meijers and Muysken (2024).

a current account surplus are reminiscent of the situation of both the Dutch and the German economy. [11]

The flows discussed above are summarized in Table 2.

Table 2 Social Accounting Matrix

	Production	Households	Firms	Foreign	Cap.Acc.	Total
Production		$C = wL + (1 - s_h)iB_F$		$p_q E - p_\mu M$	$p_q I$	$p_y Y$
Households	wL		iB_F			Y_H
Firms	π			iB_A		Y_F
Foreign						0
Savings		S_H	S_F	S_A		S
Total	$p_y Y$	Y_H	Y_F	0	$p_q I$	

Output is determined by final demand plus intermediate use (imports) and is thus equal to labor income plus the fraction of interest income that is consumed, plus investment, plus exports. Output is thus equal to:

$$Q = \left(wL + (1 - s_h)iB_{F,-1} + p_q I + p_q E\right)/p_q \tag{20}$$

As we already indicated above, we assume the price of capital goods, investment and exports equal to the output price p_q.

Real value-added or real GDP then follows from production:

$$Y = Q - M = Q(1 - \mu) \tag{21}$$

Steady state growth without inflation

The above analysis can be summarized in the savings and investment matrix – cf. Table 3. Table 3 shows that total savings in the economy S are given by savings of firms, plus the savings out of interest income by rentiers, plus savings of the foreign sector. Combining equations (14), (15), and (18), we find:

$$S \stackrel{\text{def}}{=} S_H + S_F + S_A = \pi - (1 - s_h)iB_{F,-1} - \left(p_q E - p_\mu M\right) \tag{22}$$

[11] The Netherlands had both a trade-balance surplus and a current account surplus over the whole period 1995 – 2024. The trade balance surplus fluctuated around 6 per cent of GDP in the period 1995 – 2001, then it increased steadily to a level of 11 per cent of GDP in 2014 and stabilised at that level afterwards. The current account surplus was 2,7 per cent lower on average.

To facilitate the analysis, we look at savings per lagged unit of capital σ. These are equal to:

$$\sigma = \frac{S}{p_{q,-1}K_{-1}} = \frac{\pi - (p_q E - p_\mu M) - (1 - s_h)iB_{F,-1}}{p_{q,-1}K_{-1}} \tag{23}$$

Table 3 Allocation of savings and investment matrix

	Households	Firms	Foreign	Total
Consumption	$-C = -wL - (1 - s_h)iB_F$	C		0
Investment		$p_q I$		$p_q I$
Net exports		$p_q E - p_\mu M$	$-(p_q E - p_\mu M)$	0
Wages	wL	$-wL$		0
Bonds Firms	iB_F	$-iB_F$		0
Bonds Foreign		iB_A	$-iB_A$	0
[Memo: Savings]	$[S_H]$	$[S_F]$	$[S_A]$	$[S]$
Bonds Firms	$-\Delta B_F$	$+\Delta B_F$		0
Bonds Foreign		$-\Delta B_A$	$+\Delta B_A$	0
Capital		$-p_q \Delta K$		$-p_q \Delta K$
	$\Sigma = 0$	$\Sigma = 0$	$\Sigma = 0$	$\Sigma = 0$

From equation (23) it follows that the savings per unit of capital (σ) depend positively on the profit rate $\frac{\pi}{p_{q,-1}K_{-1}}$, negatively on the trade balance through $\frac{p_q E - p_\mu M}{p_{q,-1}K_{-1}}$, and negatively on the interest rate through the term $(1 - s_h)i\frac{B_{F,-1}}{p_{q,-1}K_{-1}}$. This is plausible since a higher profit rate implies higher firm savings, which lead to higher total savings, a surplus on the trade balance increases net foreign debt and leads to more negative foreign savings, and an increase in the interest rate increases consumption of rentiers, leading to less total savings.

In Appendix A we derive that for each component does hold:

$$\frac{\pi}{p_{q,-1}K_{-1}} = (1 - \frac{p_\mu}{p_q}\mu - \frac{1}{\lambda}\frac{w}{p_q})\left(1 + \hat{p}_q\right)\left(1 + g_q\right)\frac{u_{-1}}{\kappa} \tag{24}$$

$$\frac{p_q E - p_\mu M}{p_{q,-1}K_{-1}} = \left(\epsilon - \frac{p_\mu}{p_q}\mu\right)\left(1 + \hat{p}_q\right)\left(1 + g_q\right)\frac{u_{-1}}{\kappa} \tag{25}$$

$$(1-s_h)i\frac{B_{F,-1}}{p_{q,-1}K_{-1}}=(1-s_h)ib_{F,-1}$$
$$=\left(\left(1-\epsilon-\frac{1}{\lambda}\frac{w}{p_q}\right)\left(1+g_q\right)\frac{u_{-1}}{\kappa}-g_K\right)\left(1+\hat{p}_q\right) \tag{26}$$

Substituting equations (24), (25), and (26) in (23) yields:

$$\sigma=g_K(1+\hat{p}_q) \tag{27}$$

This is consistent with equilibrium between savings and investment. The equality between savings and investment is also implied by Table 3.

In the steady state, each of these three components should be constant over time in order to constitute a constant steady state savings rate. Before we elaborate on this further, an interesting observation is that in case of steady state growth without inflation i.e. when *u*, g_K and p_q are constant over time, we find that the growth rate of capital is given by:

$$g_K^*=s_h.i \tag{28}$$

This can be explained as follows. Since the share of the capital stock financed by bonds b_F should be constant in the steady state, the capital stock should grow at the same rate as that of bonds issued by firms. We know from equations (14) and (17) that should hold:

$$S_H=\Delta B_F=s_h\ i\ B_{F,-1} \tag{29}$$

Equation (29) implies that bonds issued by firms B_F grow at a rate $s_h.i$, determined by savings behaviour of rentiers and the interest rate. This follows directly from the behaviour of households in our model: annual savings are used to buy bonds of firms and the only income from which savings are generated is interest income from rentiers. It is important to observe that this growth rate in bonds is not only restricted to a steady state but holds irrespective of the behavior of all other exogenous variables, except s_h and i.

From equations (16) and (28) we know that the utilization rate is constant in the steady state without inflation at a value:

$$u^* = \frac{1}{\beta}((\theta + s_h)\, i - \alpha) \tag{30}$$

Moreover, a constant utilization rate does imply $g_q^* = g_K^*$. Substituting this equality, together with equations (28) and (30) in equations (24) – (26) shows that each of the three components of the savings per unit of capital in equation (23) is constant over time in the steady state without inflation, hence σ is constant over time too at a value

$$\sigma^* = s_h\, i \tag{31}$$

The above holds for the steady state without inflation. The question is what happens when inflation shocks occur. From equations (24), (25) and (26) one observes that even in case of a steady inflation rate $\hat{p}_q$ the three components of savings per unit of capital σ, change over time if the utilization rate would remain constant.[12] However, since the interest rate is given, equality between savings and investment in reaction to shocks to the economy is achieved through changes in the utilization rate, which is consistent with the (post)Keynesian view. Hence the utilization rate will change in reaction to inflation. We analyze this process further in the next section.

Before we turn to the next section, we should elaborate on the development of foreign debt in the steady state. Combining equations (17) and (18) one finds

$$g_{BA} = \left(\epsilon - \frac{p_\mu}{p_q}\mu\right)\frac{p_q Q}{B_{A,-1}} + i > s_h\, i \tag{32}$$

Assuming a surplus on the trade balance, this implies that in the steady state without inflation debt of the foreign sector grows faster than output and capital. This is an obvious consequence of our assumption that the shares of both exports and imports

[12] The term $\frac{p_\mu}{p_q}\mu$ increases with constant inflation $\hat{p}_q$ – from equation (6) follows $\frac{p_\mu}{p_q}\mu = \frac{1}{(1+m)} - \frac{w}{p_q}\frac{1}{\lambda}$.

in output are constant, with $\epsilon > \mu$.[13] In that case debt held by the foreign sector increases consistently relative to output, enhanced by increasing interest payments. This debt is held by firms. Since the interest paid to firms as income of firms cancels out against the interest paid by the foreign sector as expenditures of that sector, the increasing interest payments do not affect total savings and hence have no impact on the steady state.[14]

The property of our model that in the steady state debt held by the foreign sector increases consistently relative to output is highly implausible and only made for analytical simplicity.[15] It allows us to focus on the impact of imported inflation, without complicating the model too much. To make the model more plausible, one should introduce a variable exchange rate and alleviate the assumption that exports are a constant fraction of output but depends for instance on the exchange rate and/or world trade. We leave this for further research.

The impact of inflation and variations in the markup on the utilization rate

The analysis above shows that a steady state equilibrium exists in absence of inflation. It is obvious that this equilibrium is not always stable in the presence of external shocks. We leave a derivation of the stability conditions for further research, but we observe from our simulation experiments that stability occurs under certain parameter settings. Following this observed stability in the model simulations, we assume here that the steady state equilibrium is stable.

---

[13] See footnote 11. An alternative assumption would be to choose ϵ such that for an in initial value of $\frac{B_{A,-1}}{Q}$ does hold $\epsilon = \frac{p_\mu}{p_q}\mu - \frac{1}{p_q}(1 - s_h)\frac{B_{A,-1}}{Q}$. In that case $g_{BA} = s_h.i$ would hold in the steady state without inflation and debt of the foreign sector, which might be positive or negative initially, would grow proportional to output.

[14] See also equation (22) where iB_A has dropped out.

[15] However, next to the long-lasting current account surplus, the debt of the foreign sector did fall consistently for the Netherlands for two decades by almost 200 per cent of GDP: from a positive debt position for the Netherlands of 95 per cent of GDP in 1998 to a debt by the foreign sector of 102 per cent in 2020.

As we indicated above, adjustment of the economy to external shocks occurs through changes in the utilization rate. We illustrate this for the case of an increase in import prices p_μ. Since the wage rate is assumed to be constant in our model – see also equation (6) – export prices (production prices) p_q will increase less than import prices. Consequently, the value of imports increases relative to the value of exports and a terms of trade loss occurs. This induces a decrease in profits, since the wage rate remains constant, and as a consequence output decreases – see equations (4) and (5), respectively. The decrease in output then implies a decrease in the utilization rate. Next to that, the decrease in output leads to a lower wage income, hence lower consumption and lower value-added *Y*. This enhances the decrease in output and the utilization rate falls further. The decrease in the utilization rate leads to a decrease in investment – cf. equation (16) – which consistently matches the fall in savings – cf. equation (27).

When the increase in import prices stops, the process described above reverses and the utilization rate will return to it steady state value u^*– cf. equation (30). Output growth will also return to its steady state level, cf. equation (28), but the level of output will be lower compared to the situation prior to inflation.

With respect to the adjustment process we would like to add the following observations. First, we pointed out above that those bonds issued by firms B_F grow at a rate $s_h.i$, irrespective of the behaviour of all other exogenous variables except s_h and i. Hence this process is not affected by inflation. But the growth of debt of the foreign sector B_A is affected by inflation. As we saw from equation (32), without inflation B_A grows faster than B_F does, because of the assumed trade surplus. This also holds with inflation, but then the growth rate of B_A will be lower because import prices increase faster than export prices.

Second, the profit share *h* will increase with import inflation as we explained in section 3.1 – see also equation (13). One should realize that when inflation stops, the profit share stabilizes at a higher level. The reason is that the price level of imports has increased relative to the wage rate, in which we keep the latter constant for the ease of exposition.

Third, the adjustment process illustrates the importance of distinguishing between several prices, i.e. p_μ , p_q and p_y as mentioned in section 3.1. From equations (6) and (8) one observes that at a constant rate of inflation for import prices, the output price and the GDP deflator each develop differently, and not at a constant rate. Thus, the ratios $\frac{p_\mu}{p_q}$ and $\frac{p_\mu}{p_y}$ both change over time and develop in different ways when import prices change. This affects the adjustment of the components of the savings rate to shocks in the import price – cf. equations (24) and (25). The adjustment processes are also related to the impact of a shock in import prices on the trade balance. Since output prices do not fully adjust to this shock, terms of trade losses or gains occur, and this affects output as we explained earlier in this section.

Fourth, one should realize that an increase in the output price p_q, can be caused by either an increase in import price p_μ or by an increase in the markup *m*. However, an increase in import price p_μ can lead to different consequences compared to the impact of an increase in the markup. We elaborate this point in our simulation results below.

Fifth, our assumption that the wage rate is constant has important implications for the reaction of the economy to shocks in import prices. The constant wage rate implies that output prices do not fully adjust to a shock in import prices. If the nominal wage w would be fully indexed and thus would follow the development of domestic prices, output prices would fully adjust to a shock in import prices as follows from equation (6). Then an increase in import prices would not affect output because there would be no losses in the terms of trade. Also the profit share would not be affected. This suggests that a full wage indexation to the output price is warranted to prevent rising profit shares. But this stretches the plausibility of our model too far. Moreover, including flexible wages also demands for assumptions regarding expectations and wage dynamics. It is obvious that the assumption that exports are a constant share of output, irrespective of the development of the output price, drives the absence of terms of trade losses with fully indexed wages. But that assumption is highly implausible and would require real exchange rate dynamics and assumption regarding world trade growth. This is beyond the scope of this chapter and we only mention this point here to illustrate the importance of the assumption of constant wages.

Model simulations with a variable markup

This section presents the simulation results of the model. First, we discuss the baseline outcomes without inflation. Next, we examine the outcomes of the model with import inflation and a constant markup. Afterwards, we analyze the results of the model where the markup is positively related to the rate of capacity utilization. Incorporating a variable markup precludes an analytical solution. Therefore, we rely on model simulations to illustrate the characteristics of the extended model. Finally, we show that increased import prices and a constant markup can be hardly separated from increasing the markup with constant import prices. That is, import inflation and profit inflation can be hardly separated.

The main finding is that, in the model with a fixed markup, the profit share in GDP increases if import prices rise. This mimics the result that is obtained from the analytical solutions. However, this increase in the profit share is mitigated when the markup is a positive function of the rate of capacity utilization, since the latter drops because of losses in the terms of trade as we have shown above. A variable markup thus reduces the impact of rising import prices on the profit share in GDP. Nevertheless, even with a varying markup, inflation rises, and the profit share in GDP eventually converges to the same higher level as observed under the constant markup assumption. This is due to the convergence of the rate of capacity utilization. The higher profit share persists because import prices remain higher relative to domestic prices.

Baseline results

For the baseline simulation, the initial values of the prices p_q and p_μ are set to 1. Given a value of the markup m of 0.25, an import ratio μ of 0.15, and a labour productivity λ of 2, the wage rate w is set according to equation (6). The amount of firm bonds B_F relative to nominal output remains constant in the model without inflation, consistent with equation (26). However, the amount of foreign bonds held by firms B_A relative to nominal output might remain constant, increase or decrease over time – cf. equation (32).

We assume a small positive trade balance: the export ratio ϵ of 0.2 exceeds the import ratio of 0.15. Given an imposed value for bonds, the capital stock is determined using equation (26), using $g_q = g_K^*$ from equation (28). The latter ensures that the model begins in a steady state.

The baseline simulation, which excludes imported inflation, generates a growth rate of real output equal to $g = s_h i$ – cf. equation (28). With parameter values of $s_h = \mathbf{0.8}$, and $i = \mathbf{0.01}$, this results in a growth rate of real output of 0.8%. These values are selected for illustrative purposes; while alternative parameters would affect the magnitude of the model's responses, they would not alter its fundamental mechanisms. The domestic sector has positive foreign direct investment (FDI) and together with a positive trade balance this leads to a flow of income from abroad to the domestic economy. This implies a positive current account with a growth rate of about 1%.

The output multiplier follows directly from the model presented in the previous section and can be derived by combining the production function with the definition of output – cf. equations (2) and (4):

$$p_q Q = wL + (1 - s_h) i B_{F,-1} + p_q E + p_q I \tag{33}$$

Assuming $p_q = 1$, and using the definitions of labour productivity λ, the rate of capacity utilisation u, the export ratio ϵ, and using the investment function (16), this equation can be rewritten as:

$$Q = \frac{1}{\left(1 - \frac{w}{\lambda} - \epsilon\right)} \left((1 - s_h) i B_{F,-1} + (\alpha - \theta i + \beta u_{-1}) K_{-1}\right) \tag{34}$$

This equation represents the multiplier effect of gross output arising from initial demand, which is predetermined by the previous period. The first term on the right-hand side of (34) represents the multiplier, while the second term reflects the initial final demand, which originates from households (as indicated by the first part) and investment demand from firms (in the second part). Both demand components are based on lagged variables and fixed parameters.

Given the parameter values as chosen for the baseline simulation, the value of the multiplier is approximately 6.7. Note that this value is also determined by the export ratio which is assumed to be a constant fraction of gross production.

Considering the investment function (16) we use the parameter values $\beta = 0.1$ and $\theta = 0.2$, in line with empirical findings for the Netherlands, see Meijers and Muysken (2024).[16] That is, we assume that a 1 percentage point increase of the utilization rate increases the growth rate of the capital stock by 0.1 percentage points, whereas a 1 percentage point increase of the real interest rate reduces this growth rate by 2 percentage points. Finally, we choose a rate of capacity utilization of 0.90 in the steady state, combined with a capital-output ratio κ of 1. From equation (30) it follows that the constant term in the investment function, α, is equal to -0.08. As mentioned above, the resulting growth rate of the economy is equal to 0.8% in the baseline simulation.

Turning to the initial values of the stocks, equation (26) determines the ratio between B_F and $p_k K$. Given the selected parameters and assuming no inflation this bonds-to-

[16] Our simulation results did show that for higher values of β the stability of the model is affected.

capital-ratio is equal to about 64. This value will be lower if we introduce autonomous final demand, for instance by government or by households. We select values of $B_F = 10$ and $B_A = 5$ as starting values.

All the above results in a baseline profit share in GDP of 22.2%. The trade balance surplus is 5% of GDP and the interest inflow $i.B_A$ is also positive, leading to a growth of the current account surplus to GDP ratio. Inflation is zero in the baseline scenario and the rate of capacity utilization is 0.9 as determined by the parameters.[17]

Impulse to import prices with a constant markup

In the case of a fixed markup, the impact of an increase in import prices on the profit rate has already been elaborated analytically above. Here, we briefly present the numerical simulation results. We apply two scenarios: one in which import prices increase by 10% per period during 5 periods (Scenario 1), and Scenario 2 in which import prices increase by 5% per period during 20 periods. Scenario 1 examines the effects of a large but short-term increase in import inflation, while Scenario 2 explores the effects of a more prolonged, gradual increase. Figures 1 and 2 show the results of these simulations.

The left-hand panel of Figure 1 shows the results of the impulse in import prices. In the baseline scenario, the growth rate of import prices is zero; in scenario 1, it is 10% during 5 periods; and in scenario 2, it is 5% during 20 periods. Note that the resulting growth rate of output prices is not constant. As follows from equations (6) and (7), with a constant wage rate and constant labor productivity, the growth rate of the GDP deflator increases even when import inflation is constant. The growth rate of output prices also varies over time, as follows from equation (8).

As follows from equation (12), the profit share in GDP rises when import prices increase. It stabilizes at a higher level because of import prices are permanently higher

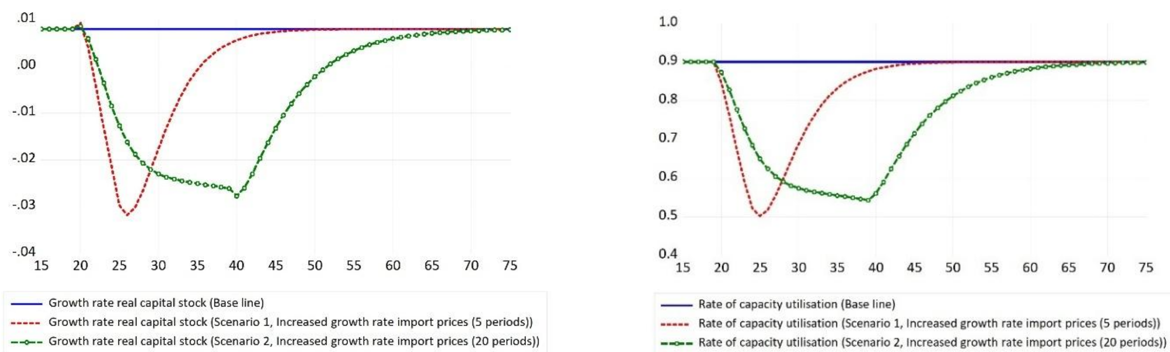
[17] The model code and the parameter values are included in Appendix B.

after the inflationary episode. This result is replicated in the numerical simulation and is displayed in the right-hand panel of Figure 1.

Figure 1: Simulations

In the first period following the impulse, inflation begins to rise, leading to a decrease in the real interest rate, given a constant nominal rate. The growth rate of the real capital stock then increases, as shown in the left-hand panel of Figure 2. However, it soon drops significantly since the rate of capacity utilization starts to decline immediately after the impulse, as depicted in the right-hand panel of Figure 2. In subsequent periods, the rate of capacity utilization drops further and soon outweighs the positive effects of reduced investment costs, causing the growth rate of the capital stock to decrease.

Figure 2: Simulations



After the impulse to import inflation has ended, an opposite shock occurs: the growth rate of the real capital stock declines as inflation abruptly drops to zero. Subsequently, both the growth rate of the real capital stock and the rate of capacity utilization gradually return to their initial levels, albeit at a very slow pace.

Impulse to import prices with a variable markup

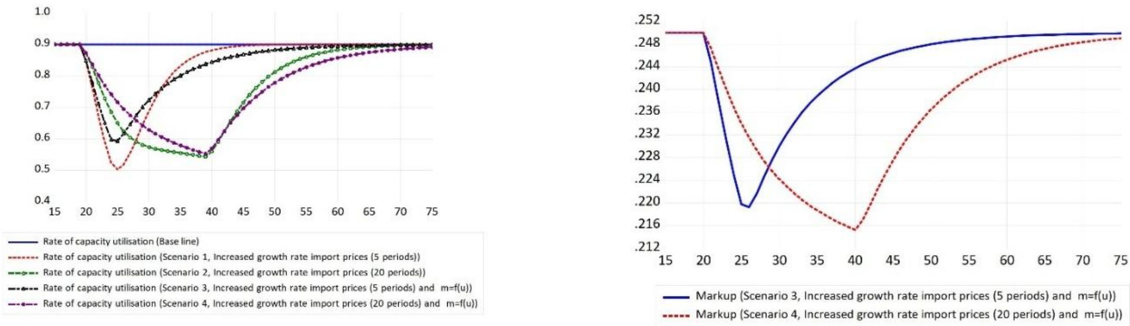
The main purpose of simulating the model is to incorporate a variable markup. As explained above, the results of such a model cannot be solved analytically. Following Hein (2014, p 189) we assume a positive relation between the markup and the rate of capacity utilization. This stems from the notion that unit overhead costs, or unit fixed costs, increase at output levels above some normal rate of capacity utilization as a result of which firms increase the markup to cover these fixed costs. For simplicity, we assume that the markup is a linear function of the rate of capacity utilization:

$$m = m_0 + m_1 . u_{-1} \tag{35}$$

We choose a value of 0.1 for m_1 and since the rate of capacity utilization is equal to 0.90 in the steady state, m_0 is 0.16 to maintain the same value for the markup as in the baseline simulation. The same impulses to import prices as used earlier are applied here, and the baseline simulations are the same as before. Scenarios 3 and 4 are comparable to Scenarios 1 and 2, but they now incorporate a variable markup. The results of these simulations are shown in Figures 3, 4 and 5.

The left-hand panel in Figure 3 shows that the rate of capacity utilization is slightly less responsive to an increase in import price inflation in case of a variable markup, though it still decreases considerably. This indeed implies that the markup itself decreases as is depicted in the right-hand panel of this figure.

Figure 3: Simulations

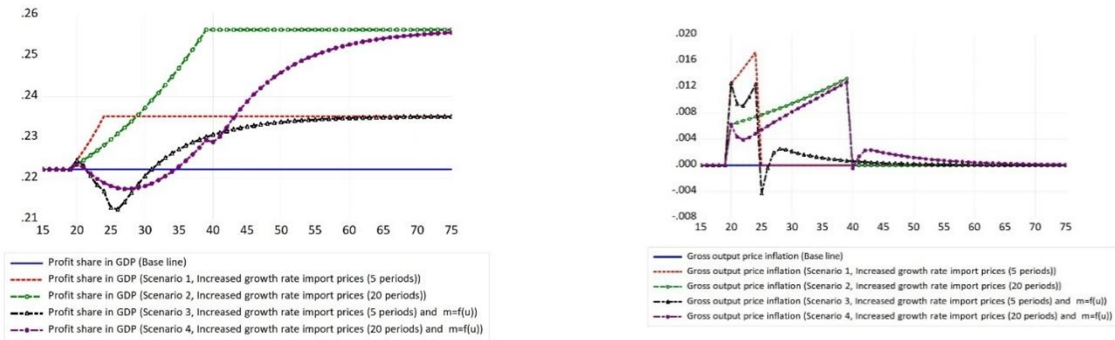


In case of a constant markup, the profit share increases immediately following inflation in import prices as we observed in scenarios 1 and 2. However, in case of a variable markup, the decreasing markup implies that the profit share in GDP will decline initially. This is illustrated in the left-hand panel of Figure 4.

To understand this process, one should realize that the profit share is positively dependent on the ratio of the markup and the wage share in total costs, m/a – see also equation (11). A higher markup leads to a higher profit share as we argue in section 3.2. As long as the markup does not change, a higher wage share will by definition lead to a lower profit share in total costs and also a lower profit share h in value added. If we apply these insights to the present analysis, we know that during the impulse, p_μ increases. As a result of the declining rate of capacity utilization, the markup decreases. However, the labor share in total costs a also decreases due to rising import prices. Initially, the effect of the declining markup dominates, leading to a decrease in the profit share. As p_μ continues to increase, the labour share in total costs continues to fall, while the rate at which capacity utilisation declines gradually slows. Consequently, the pace at which the markup decreases also slows. Eventually, the impact of the declining labour share in total costs outweighs that of the falling markup, causing the profit share in GDP to rise and ultimately exceed the baseline level with constant import prices.

When inflation stops in period 2039, the profit share slowly converges to the steady state value that is also obtained in the case of fixed markup.

Figure 4: Simulations



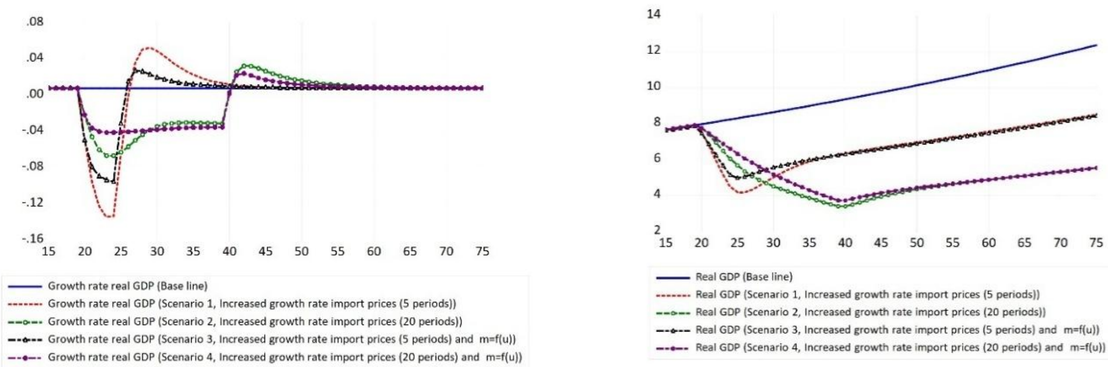
Focusing on the rate of inflation of domestic prices, the right-hand panel of Figure 4 shows that a 10% (5%) increase in import inflation leads to about 1.7% (1.3%) increase of domestic inflation. In case of a varying markup the domestic inflation is lower than in case of a fixed markup but eventually converges to the same rate. This process is sluggish, however, such that the rate of domestic inflation remains lower for a considerable period.

Concentrating on the development of real GDP in these four scenarios, the GDP growth rate becomes negative following a shock in import prices. Due to the loss in terms of trade, domestic income decreases, leading to a decline in GDP. The fixed export ratio amplifies this effect. The GDP growth rate eventually stabilizes, as is more clearly seen in scenarios 2 and 4 in the left-hand panel of Figure 5, where import inflation persists for a longer period as compared to scenarios 1 and 3.

In case of a fixed markup, the GDP growth rate declines more sharply compared to a scenario where the markup varies with the rate of capacity utilization. This difference is driven by variations in labor income across the two scenarios. As observed earlier, the profit share in GDP is lower when the markup is variable, implying a higher labor share. Since the marginal propensity to consume out of labor income is one, a higher labor income share results in greater domestic demand.

In the long run, the GDP growth rates all converge. Although the GDP growth rate eventually returns to the baseline level, the level of real GDP remains lower after the price shock, as shown in the right-hand panel of Figure 5. This is due to permanently higher import prices relative to domestic prices, leading to sustained losses in terms of trade.

Figure 5: Simulations

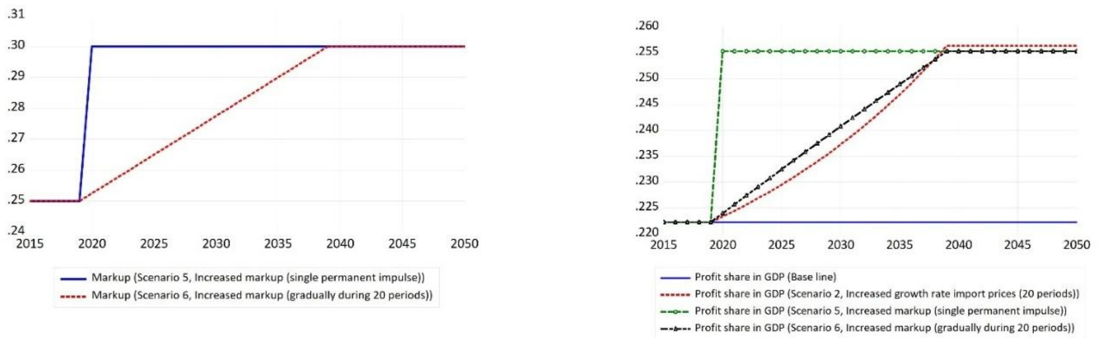


The impact of an increased markup

Increasing the markup and raising import prices can have comparable effects on the economy. This is examined in scenarios 5 and 6, where we compare the effects of higher import inflation with those of an increased markup. Figures 6 and 7 show the results of this experiment.

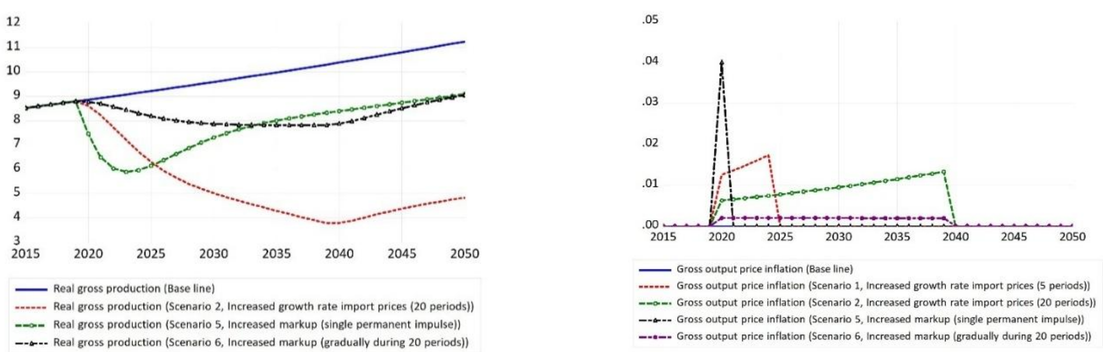
In scenario 5, the markup is increased from 0.25 to 0.30 in a single period, whereas in scenario 6, the markup gradually rises over 20 periods from its initial value of 0.25 to 0.30 (see the left-hand panel of Figure 6). The right-hand panel shows the effects on the profit share in GDP, which, in the case of a gradual markup increase, are comparable to the profit share in scenario 2.[18] However, the effect on output differs as we elaborate below.

Figure 6: Simulations



An increase in import prices leads to a decline in the terms of trade, reducing both output and GDP. In contrast, a rise in the markup alters the distribution of GDP between profit and wage shares but leaves the terms of trade unchanged. As a result, the reduction in real output and GDP is smaller, as can be seen in the left-hand panel of Figure 7.

Figure 7: Simulations



The impact on inflation also differs. While the effect of a higher markup on the profit share in GDP is similar to that of rising import prices, its influence on inflation is

[18] We ignore the comparison with scenario 3, to keep the Figures transparent.

distinct. A one-time increase in the markup causes a temporary inflation spike, as shown in the right-hand panel of Figure 7. In contrast, a gradual increase in the markup leads to only a slight rise in inflation—significantly lower than the inflation caused by import price increases, despite the comparable effects on the profit share.

Conclusions from the simulations

The simulation results examine the numerical effects of import inflation on domestic inflation, the income distribution, and GDP. We focus on variations in the markup which cannot be analyzed analytically. In the baseline scenario, where inflation is absent, the economy grows steadily at 0.8%, with a profit share of 22.2% and a positive trade surplus.

A second scenario introduces a variable markup, which adjusts to the rate of capacity utilization. The results show that allowing the markup to fluctuate, mitigates the immediate impact of rising import prices on the profit share in GDP. While GDP growth still declines in response to higher import prices, the contraction is less severe compared to the fixed markup scenario. Inflation also rises, but to a lesser extent than when the markup remains constant. In the longer run the inflation rates converge to the scenario with a fixed markup though this process is slow. Another interesting difference with the fixed markup is that the profit share initially decreases in response to imported inflation, whereas it immediately increases in case of a fixed markup.

In the last scenario we compare the effects of import inflation with those of a direct increase in the markup. The results reveal that raising the markup and increasing import prices can have comparable effects on the economy, particularly in terms of increasing the profit share in GDP. However, an increase in markup does not affect the terms of trade, meaning that the reduction in real GDP is smaller than in the case of import inflation. Additionally, a sudden increase in the markup leads to a one-period spike in inflation, whereas a gradual increase results in a more moderate inflationary effect, compared to the increase of import prices.

Summary and concluding remarks

This chapter analyses the relation between import inflation, the profit share, and markups. The context is the discussion about profit inflation that has occurred because of rising import prices. We first show that in case of a constant markup, the profit share can increase if import prices increase. However, an increase in the profit share following an increase in import prices does not necessarily imply a change in the markup behavior of firms. To elaborate this point, we develop a three-sector model for an open economy to further analyze the consequences of an increase in import prices on the profit share. In this model, for a given interest rate, inflationary shocks are absorbed by a change in the utilization rate. We first analyze in our model the case of a constant markup. We then simulate the model for the case of a variable markup, by linking the markup to the utilization rate. We simulate the model for 2 scenarios: Scenario 1 examines the effects of a large but short-term increase in import prices, while Scenario 2 explores the effects of a more prolonged, gradual increase in import prices. The simulation results of this model show that in case of a fixed markup, an increase in import prices leads to an increase in the profit share and to a permanent higher level of this share. In addition, at first the utilization rate and the capital stock decrease but return to their initial levels as the increase in import prices comes to an end. The adjustment period is longer in scenario 2 than in scenario 1. If we link the markup positively to the utilization rate, we find similar patterns for the utilization rate and given the link between the utilization rate and the markup, the markup first decreases and thereafter increases. This adjustment is reflected in the profit share, which first decreases and thereafter increases. When the inflation stops, the profit share slowly converges to the steady state value that is also obtained in the case of a fixed markup. In our model a 10% (5%) increase in import inflation leads to about 1.7% (1.3%) increase of domestic inflation. In case of a variable markup the domestic inflation is lower than in case of a fixed markup but eventually converges to the same rate. Finally, we compare the effects of import inflation with those of a direct increase in the markup. The results reveal that raising the markup and increasing import prices can have comparable effects on the economy, particularly in terms of increasing the profit share in GDP. From this we conclude that without further research, it is very difficult to ascribe inflation that coincides with an increase in import prices to either the change in import prices or to an increase in the markup.

Our analyses have important policy implications with respect to the profit share. We show that in case of increased import prices, inflation will occur, and the profit share will increase, irrespective whether the markup increased or not. However, when the markup increases, the profit share will increase stronger. In that sense it is interesting to observe that when the markup is positively related to the utilization rate, an initial fall in the utilization rate after import price increases induces a lower increase or even in a decrease in the profit share. This also implies that policies aiming at maintaining a stable profit share should influence price behavior of firms to mitigate the increase in the profit share at the detriment of the wage share. Examples of such policies are competition and anti-monopoly policies, price controls and increased (excess) profit taxes.

A final observation is that in our analysis we have assumed that wages are constant. It seems plausible that in all scenarios studied above, the profit share increases at the cost of the wage share in case of incomplete wage indexation. The outcome might be different in case of complete wage indexation. However, that is a question for further research.

References

Acharya, V., Crosignani, M., Eisert, T., & Eufinger, C. (2024, February 6). The euro area pandemic inflation: A tale of supply chains, household inflation expectations, and firm pricing power. *VOX EU*.

Arce, O., Hahn, E., & Koester, G. (2023, March 30). How tit-for-tat inflation can make everyone poorer. *The ECB Blog*. European Central Bank.

Arrow, K. (1974). The measurement of real value added. In P. A. David & M. W. Reder (Eds.), *Nations and households in economic growth* (pp. 3–19). Academic Press.

Bernanke, B., & Blanchard, O. (2023). What caused the U.S. pandemic-era inflation? *NBER Working Paper* No. 31417.

Bernanke, B., & Blanchard, O. (2024). An analysis of pandemic-era inflation in 11 economies. *NBER Working Paper* No. 32532.

Brochier, L. (2020). Conflicting-claims and labour market concerns in a supermultiplier SFC model. *Metroeconomica, 71*(3), 566–603.

Capolongo, A., Kühl, M., & Skovorodov, V. (2023, May 12). Firms' profits: Cure or curse? *The ESM Blog*. European Stability Mechanism.

Colonna, F., Torrini, R., & Viviano, E. (2023). The profit share and firm mark-up: How to interpret them? *Occasional Papers No. 770*, Banca d'Italia.

Cucignatto, G., Garbellini, N., & Fora-Alcalde, F. (2023). Profit-led or cost-led inflation? Propagation effects through the EU inter-industry network. *PSL Quarterly Review, 76*(306), 261–276.

De Keyser, T., Langenus, G., & Walravens, L. (2023). The development of corporate profit margins and inflation. *NBB Economic Review*, 2023, No. 8.

Eeckhout, J. (2021). *The profit paradox*. Princeton University Press.

Erken, H., Groot, S., Schreijen, S., & van der Veen, M. (2023). Is in Nederland sprake van 'graaiflatie'? *RABO Research Publication*, May 2023.

Ferguson, T., & Storm, S. (2023). Myth and reality in the great inflation debate: Supply shocks and wealth effects in a multipolar world economy. *International Journal of Political Economy, 52*(1), 1–44.

Gallo, E. (2024, February 27). Unpacking decreasing inflation: Profit inflation vs. the heroic central bank. *Monetary Policy Institute Blog* #122.

Gerinovics, R., & Metelli, L. (2023, December 18). The evolution of firm markups in the US and implications for headline and core inflation. *VOX EU*.

Glover, A., Mustre-del-Río, J., & von Ende-Becker, A. (2023a, January 12). How much have record corporate profits contributed to recent inflation? *Economic Review*. Federal Reserve Bank of Kansas City.

Glover, A., Mustre-del-Río, J., & Nichols, J. (2023b, May 12). Corporate profits contributed a lot to inflation in 2021 but little in 2022 – A pattern seen in past economic recoveries. *Federal Reserve Bank of Kansas City Bulletin*.

Hahn, E. (2023). How have unit profits contributed to the recent strengthening of euro area domestic price pressures? *ECB Economic Bulletin, Issue 4/2023*.

Hansen, N.-J., Toscani, F., & Zhou, J. (2023). The role of import prices, profits, and wages in the current inflation episode in the euro area. *IMF Working Paper* 23/131. International Monetary Fund.

Hebbink, G., & Öztürk, B. (2023, May). De bijdrage van winsten en lonen aan de Nederlandse inflatie. *DNB Publicatie*.

Hein, E. (2006). Money, interest and capital accumulation in Karl Marx's economics: A monetary interpretation and some similarities to post-Keynesian approaches. *The European Journal of the History of Economic Thought, 13*(1), 113–140.

Hein, E. (2012). *The macroeconomics of finance-dominated capitalism – and its crisis*. Edward Elgar Publishing.

Hein, E. (2014). *Distribution and growth after Keynes: A post-Keynesian guide*. Edward Elgar Publishing.

Kalecki, M. (1971). *Selected essays on the dynamics of the capitalist economy 1933–1970*. Cambridge University Press.

Kalemli-Özcan, S., Silva, A., Yildirim, M., & di Giovanni, J. (2022). Global supply chain pressures, international trade, and inflation. *CEPR Discussion Paper* No. DP17449.

Konczal, M., & Lusiani, N. (2022, June 21). Prices, profits, and power: An analysis of 2021 firm-level markups. *Roosevelt Institute Brief*.

Lavoie, M. (2022). *Post-Keynesian economics: New foundations*. Edward Elgar Publishing.

Lavoie, M. (2023, May 14). Some controversies in the causes of post-pandemic inflation. *Monetary Policy Institute Blog* #77.

Meijers, H., & Muysken, J. (2024). A post-mortem of interest rate policy: Beware of financial fragility. *UNU-MERIT Working Paper* 2024-030, UNU-MERIT, Maastricht.

Meijers, H., Muysken, J., & van Veen, T. (2025). Rentebeleid geen oplossing voor gevolgen inflatie. *Economisch Statistische Berichten*. 110 (4842), 63-65.

Nikiforos, M., Grothe, S., & Weber, J. D. (2024). Markups, profit shares, and cost-push-profit-led inflation. *Industrial and Corporate Change, 33*(2), 342–362.

Pressman, S. (2023, May 18). Covid and the causes of inflation. *Monetary Policy Institute Blog* #79.

Rudd, J. (2021). Why do we think that inflation expectations matter for inflation? (And should we?). *Finance and Economics Discussion Series* 2021-062. Board of Governors of the Federal Reserve System.

Setterfield, M. (2022). Inflation and distribution during the post-COVID recovery: A Kaleckian approach. *WP 09/2022*, Department of Economics, The New School for Social Research.

Sims, C. (1969). Theoretical basis for a double deflated index of real value added. *The Review of Economics and Statistics, 51*(4), 470–471.

Soederhuizen, B., Bettendorf, L., & Meijerink, G. (2023, October). Uitsplitsing van de inflatie in Nederland. *CPB Publicatie*, The Hague, The Netherlands.

Storm, S. (2022). Inflation in a time of corona and war. *INET Working Paper No. 185*. Institute for New Economic Thinking.

Storm, S. (2023, June 15). Profit inflation is real. *INET*. Institute for New Economic Thinking.

Storm, S. (2024). Tilting at windmills: Bernanke and Blanchard's obsession with the wage-price spiral. *International Journal of Political Economy*, *53*(2), 126–148.

Appendix A Derivation of equations (24) – (26)

Starting from

$$\sigma = \frac{S}{p_{q,-1}.K_{-1}} = \frac{\pi - \left(p_q E - p_\mu M\right) - (1 - s_h) i.B_{F,-1}}{p_{q,-1}.K_{-1}} \tag{23}$$

We first point out that does hold:

$$\frac{p_q Q}{p_{q,-1}.K_{-1}} = \left(1 + \hat{p}_q\right)\left(1 + g_q\right)\frac{u_{-1}}{\kappa} \tag{A1}$$

using equations (1), (3) and the definitions of the growth rates of p_q and *Q*.

Second, we find:

$$\pi = p_q Q - p_\mu M - wL = (1 - \frac{p_\mu}{p_q}\mu - \frac{1}{\lambda}\frac{w}{p_q})p_q Q \tag{A2}$$

using equations (5) and (19). Combining (A1) and A2) then yields:

$$\frac{\pi}{p_{q,-1}.K_{-1}} = (1 - \frac{p_\mu}{p_q}\mu - \frac{1}{\lambda}\frac{w}{p_q})\left(1 + \hat{p}_q\right)\left(1 + g_q\right)\frac{u_{-1}}{\kappa} \tag{24}$$

Third combining equation (19) and (A1) leads to:

$$\frac{p_q E - p_\mu M}{p_{q,-1}.K_{-1}} = \left(\epsilon - \frac{p_\mu}{p_q}\mu\right)\left(1 + \hat{p}_q\right)\left(1 + g_q\right)\frac{u_{-1}}{\kappa} \tag{25}$$

Finally, we know from equation (15) and Table 3:

$$S_F \ = \ \pi + i.B_{A,-1} - i.B_{F,-1} = \ \Delta B_A + p_q.\Delta K - \Delta B_F \tag{A3}$$

Substituting equation (17) for $\Delta B_A = -S_A$ and substituting (18)

$$\pi + i.B_{A,-1} - i.B_{F,-1} = \left(p_q E - p_\mu M\right) + iB_{A,-1} + p_q.\Delta K - \Delta B_F \tag{A4}$$

Using equations (19), (29) and (A1) we then find:

$$\left(1 - \frac{w}{p_q\lambda} - \epsilon\right)\frac{u}{\kappa}(1 + g_q)\left(1 + \hat{p}_q\right)p_{q,-1}.K_{-1} \tag{A5}$$

$$= g_K.\left(1 + \hat{p}_q\right)p_{q,-1}K_{-1} + (1 - s_h)iB_{F,-1}$$

Rearranging equation (A5) yields:

$$(1-s_h)i\frac{B_{F,-1}}{p_{q,-1}.K_{-1}} = (1-s_h)i.b_{F,-1} \\ = \left(\left(1-\epsilon-\frac{1}{\lambda}\frac{w}{p_q}\right)(1+g_q)\frac{u_{-1}}{\kappa}-g_K\right)(1+\hat{p}_q) \tag{26}$$

Appendix B Model code and parameter values[19]

pmu = pmu(-1) * (1 + par_gpmu + add_gpmu)	'price level import prices
pimu = @pch(pmu)	'growth rate import prices
markup = par_m1 + par_m2*u(-1)	'markup
pi = pq * Q - pmu * M - W * L	'profits
pq = (1 + markup) * (w / par_lambda + pmu * par_mu)	'profits
pq = (1 + markup) * (w / par_lambda + pmu * par_mu)	'price level gross output
piq = @pch(pq)	'growth rate gross output prices (inflation)
py = (pq - pmu * par_mu) / (1 - par_mu)	'price level GDP deflator
piy = @pch(py)	'growth rate gdp prices (GDP deflator)
h = pi / (py * Y)	'profit share in GDP
L = Q / par_lambda	'employment
M = par_mu * Q	'real imports
E = par_eps * Q	'real exports
Yh = w * L + par_i * BF(-1)	'income households
Cons = (1-par_sw) * w * L + (1 - par_sh) * par_i * BF(-1)	'nominal consumption
Sh = Yh - Cons	'savings of households
Yf = pi + par_i * BA(-1)	'income of firms
Sf = Yf - par_i * BF(-1)	'savings of firms
Y = (Cons + pq * E - pmu * M + pq * I) / py	'Real GDP
Q = Y + M	'Real gross output
Sa = - (pq * E - pmu * M) - par_i * BA(-1)	'Savings foreign sector
S = Sh + Sf + Sa	'total savings
Ba = Ba(-1) - Sa	'Bonds abroad
Bf = BF(-1) + Sh	'Bonds firms
gk = par_alpha+par_beta*u(-1)-par_theta*(par_i -piq)/(1+piq)	'growth rate capital stock
I = gk * K(-1)	'Investment in fixed capital
K = K(-1) + I	'Capital stock
Qc = K / par_kappa	'Full capacity gross output
u = Q / Qc	'rate of capacity utilization
w = w(-1) * (1 + par_gw)	'nominal wage rate
qn = q * pq	'nominal gross output
yn = y * py	'nominal GDP
kn = k * pq	'nominal capital stock
Wf = Kn+Ba-Bf	'net wealth of firms
Wh = Bf	'net wealth of households
Wa = -Ba	'net wealth of foreign sector
tradeb = pq*E-pmu*M	'trade balance
CA = tradeb + par_i*BA(-1)	'Current account
CP = d(BA)	'Capital account
BP = CA - CP	'Balance of payment

[19] The prescript “par_” refers to a parameter and “add_” to an add factor that is used in the scenarios. (@pch() calculates the growth rate, and d() represents the first difference)

For the base line simulation, we checked that $S = p_q I$ which is indeed the case.

Parameter values:

par_lambda = 2	'labor productivity
par_kappa = 1.0	'capital-output ratio
par_gw = 0	'growth rate nominal wages
par_gpmu = 0	'growth rate import prices
par_beta = 0.1	'sensitivity of investments to the rate of capacity utilisation
par_theta = 0.2	'sensitivity of investments to the real interest rate
par_eps = 0.15	'exports/gross output ratio
par_mu = 0.1	'imports/gross output ratio
par_i = 0.01	'nominal interest rate
par_sh = 0.8	'marginal consumption out of capital income
par_sw = 0.0	'marginal consumption out of wage income
par_m1 = 0.25	'markup in scenarios 3 to 5 reduced to 0.16
par_m2 = 0.00	'effect of u on markup, in scenarios 3 to 5 increased to 0.1

Calculated Initial values and remaining parameters:

u = 0.9	'initial rate of capacity utilization
bf=10	'initial bonds held by households
ba=5	'initial bonds held by firms
w=pq/(1+markup)-pmu*par_mu)*par_lambda	'wage rate, initial value is 1.4
par_alpha par_alpha= par_i*par_sh-par_beta*u+par_theta*par_i	'constant in investment function
K = (Bf/pq)/(((1-w/(pq*par_lambda)-par_eps)*(1+gk)*u/par_kappa - gk)*(1+piq)/((1-par_sh)*par_i))	'initial capital stock
piq = 0	'initial output price inflation
gk = par_alpha + par_beta * u - par_theta * (par_i - piq)/(	'initial growth rate capital stock
Qc=K / par_kappa	'initial potential output
Q = Qc* u	'initial actual output
py = (pq-pmu*par_mu)/(1-par_mu)	'initial GDP price level

Asset prices and exchange rates: a time dependent approach

Giulia Piccillo

Abstract

This study explores the connection between exchange rates and asset prices using an order flow framework, emphasizing the effects of time-varying risk aversion. The model captures regime-switching behavior, alternating between optimistic and pessimistic equilibria. Empirical analysis on the UK, Switzerland, and Germany (with Japan as a benchmark) confirms the presence and persistence of these regimes. The results are consistent with literature on investor risk preferences and provide novel insights, especially in the context of financial crises. Notably, three out of four markets conform to both order flow dynamics and Markov-switching models, offering a refined approach to exchange rate modeling.

Introduction

The relationship between the exchange rates and the stock prices has been a rich field of study for several decades. This is because we observe that the two markets have high (and similar) volatilities that are not justified by a purely fundamental analysis. Several papers have been written describing the statistical relationship between the two markets. However not always a solely statistical study may uncover the connection between the two (Hatemi-Ja and Roca, 2005; Hatemi-Ja and Irandoust, 2002; Ramasamy and Yeung, 2005; Granger et al., 1998). A theory drawing the underlying picture is at the core of most of the contributions. To this day there is no commonly accepted framework in the literature. The focus of the most recent research is the theory to choose in order to trace this relationship.

The papers on the subject may be divided in two classes. One is macro, observing the flows as a result of large economic fundamentals. The second one is micro, related to financial measures and how they affect prices and profit maximizing rules. Both make strong cases and are in fact probably working together in the determination of the exchange rate and asset price time series.

This paper will take a micro approach. Specifically it fits into the set of models describing the microstructure of exchange rates. The theoretical intuition comes from papers on capital flows. The application to exchange rate literature has been recent (Evans and Lyons, 2002). Since then it has quickly spread due to the simplicity of the concepts as well as to the strong empirical results obtained (Evans and Lyons, 2006). More recently one of the papers on micro-structure has drawn a complete picture of the exchange rate through order flows and has tested the implications of the model on commonly available data. The great majority of the markets studied is consistent with the theory. The paper was published in the Review of Financial Studies by Hau and Rey (2006). Our paper is adopting the Hau-Rey framework.

Another stream of literature that has proven successful in increasing the forecasting power of previous models has been the time-dependent approach. There are several kinds of time-varying statistical models. The most flexible is typically GARCH. Then comes the literature on structural breaks. Finally, Markov Switching Models (MSM) are among the most regulated time-varying frameworks. In this paper we use these last set of models, due to the strong structure that it holds in the estimation (Otranto and Gallo, 2002; Alvarez-Plata and Schrooten, May 2003; Sims and Zha, 2004; Cheung and Erlandsson, December 2004; Psaradakis and Spagnolo, 2003; Bellone, 2005; Krolzig, 1996).

The main trait of this kind of models is that it alternates between a limited number of specifications. This means that the program is not allowed to create a better fitting model every time there is a break. Instead it is required to look for similarities throughout time. This is conceptually very similar to a representative agent's behavior, who looks at the past to interpret the present and forecast the future. In order to impose these preferences, a MSM is then the most desirable one. The two streams of literature are merged in the model devised in the following paragraphs. Specifically, sections 2 and 3 show an order flow model that is subject to switching regimes. Section 4 carries out an empirical analysis to verify whether the theory is consistent with the data. We have also thought it interesting to describe how the results obtained by the model may be interpreted at the light of recent events and literature on the subject. Finally a short conclusion will summarize and highlight future research possibilities.

The order flow model

This paper is based on the theoretical framework described in Hau and Rey (2006). It focuses on the dynamics of the exchange rate in relation with the stock prices and capital flows. The value taken by the exchange rate is a function of the order flows originating from the capital market. This may be referred to as the micro-finance approach to exchange rates.

Intuition

There are two countries, Home and Foreign. Each country has two assets: stocks and bonds. This means that in the world there are four different assets. However, only the stocks are traded internationally. Investors are not allowed to hedge their exchange rate risk when investing in the capital market. As a result only the stock market dynamics affects the evolution of the exchange rate.

The exchange rate takes on the value where the demand for foreign currency meets the supply. The foreign trade in capital assets necessarily affects the demand for foreign currency. Specifically, two major transactions cause the demand to rise. The first is given by new local investors wanting to buy foreign shares, the second is due to local companies giving out dividends that are going to be returned to the foreign investors. Both of these actions cause an increase in the demand for foreign currency balanced by a sale of local currency. In a symmetric way foreign investors buying local company shares and dividends given to local investors by foreign companies will contribute to the Home currency demand.

Facing this demand there is a supply mechanism. Foreign currency is supplied by institutions that give out liquidity with a certain elasticity (i.e., k). So the liquidity is given out as the pressure on the price of the foreign currency (exchange rate) gets heavier. This means that the elasticity has an important role in the determination of how much liquidity is supplied to the market, and therefore in setting the exchange rate.

Given the importance of the elasticity of liquidity supply we will describe it further. We will consider a world where the exchange rate follows a Ornstein-Uhlenbeck process that reverts to a stable equilibrium S at a constant speed αe. The investors in foreign currency maximize a standard mean variance objective. Solving the model, the value of k is a function of the degree of risk aversion and of the statistical characteristics governing the O-U process (namely the speed of convergence to equilibrium and the exchange rate variance).

This means that a change in elasticity is caused by either a change in risk aversion or a change in the exogenous statistical properties of the exchange rate. More intuitively, if risk aversion decreases the investors are going to be happy with a lower profit for the same variance of the exchange rate. If the equilibrium value of the exchange rate does not change, the excess supply of foreign currency will decrease through a lower elasticity. In the next section, we will make the assumption, following a branch of behavioral finance literature, that the degree of risk aversion will change over time. We will derive what should happen if it changes and compare the theoretical implications with the data.

Back to the description of the model, the dynamics of the stock prices is studied. These prices are a function of the local dividends. The steady states of the dividends and their expected values are also important in the determination of the value of the stocks. The dividends also follow a O-U process.

The model solves for a unique solution under market incompleteness (with respect to foreign exchange risk). This solution becomes equivalent to the one in a complete market in the limit where the currency elasticity of supply is infinite. Therefore, the final solution is also a function of the value of k (for a formal description of the model please refer to the online appendix).

Implications

The theory described has several important implications that allow its testing in an intuitive manner. The first result is that (for certain parameters) exchange rate returns have almost as much volatility as equity returns. The next result is the most relevant. It may be shown that foreign stock returns and exchange rate returns are negatively correlated.[1] In other words a positive foreign stock return will cause a depreciation of the foreign currency.

The intuition behind this is that equity trade is the source of order flows: a positive foreign dividend innovation will create a sequence of consequences. It increases the dividend of foreign stock. An increase in the dividend has two effects. The first one is that the price of the stock rises, causing the returns to increase as well. The second one is that the demand for foreign currency decreases. This is mainly due to the repatriation of the dividends (of the home investors). Foreign currency is sold versus the home currency, causing the foreign currency to depreciate.[2]

Following this implication the exchange rate is an automatic partial hedge against foreign stock risk. With a similar intuition, the relative stock returns have an exactly inverse correlation with the returns of the exchange rates.[3] Empirically a simple negative correlation will be considered sufficient due to the presence of exogenous shocks as well as of country asymmetries.

---

[1] For the sake of synthesis, I will spare the formal proofs of these implications, that may be found in Hau and Rey, 2006.

[2] There is also another reason why the foreign currency is sold. It may be called the risk re-balancing channel after the paper by Hau and Rey. It means that the home investor will reduce its holding of foreign stock because of the larger exposure from the increased value of its shares.

[3] The relative stock returns are defined as simply the difference between the local and the foreign stock return at every time period.

Finally, the authors interpret the change of their empirical data over time (and in cross-sections) as a function of equity market development. So the negative correlation between the two markets increases as the equity market is more developed. While the article shows other implications, the ones described above are enough for our discussion.

What if parameters change over time?

The contribution of this paper is to assume that the economy changes over time within the framework offered by Hau and Rey's model. Specifically, we will formulate the assumption that one of the exogenous parameters evolves according to a Markov chain. Of the "constants" considered in the model there is one that is commonly considered to be non-varying in a shorter period. This is risk aversion. On the one hand, to consider this as a constant in the long run would mean to assume that the overall perception of risk stays similar throughout time. On the other hand, it would be possible to make the opposite hypothesis that this is always changing according to no scheme at all. This would imply the forceful assumption that the agents' understanding, and interpretation of the world keeps adjusting in manners always new and not based on past schemes.

Therefore, it is natural to assume for the time being that the parameter for risk aversion is indeed time-dependent, but within a certain scheme. This means that agents recognize the current situation as similar (under the risk approach point of view) to other past occurrences. Therefore, they will behave accordingly, in a similar fashion to how they did in the past. In the remainder of this paper, we will also assume that the other exogenous statistic properties of the exchange rate remain stable (in mean) through time. The motivation for this assumption is that the order flow model already has rules on the evolution of the exchange rate. To modify them and to assume that they change exogenously would be to forcefully add explanatory power without adding a sound economic reasoning. The focus of this paper is to integrate an assumption that is popular in the financial literature with a promising new scheme for the foreign exchange market.

Changing equilibrium

It is reasonable to assume that the agents behave according to two separate rules. In other words, we are going to assume that there are two different (and alternating) approaches to risk. Obviously, one will be more optimistic, while the other will be more pessimistic. This type of agent behavior may be statistically described by a Markov Switching model. Both values taken by the parameter for risk aversion ρ will be within the limits of the theory described above.[4] By this we assume that in both cases there will exist one and one only equilibrium of the system.

This is defined as the unique equilibrium of the relationship between the exchange rate and the stock prices of the two markets. While the theory concentrates on the definition of the steady states of the three variables (the exchange rate and the two countries' stock prices) we will focus our analysis on the coefficient (also affected by the change in risk aversion) that relates the first difference of the exchange rate to the relative stock returns. Since risk aversion varies according to regimes also the equilibrium of the system (which is function of it) will shift regime when the risk aversion parameter changes. This is as from Hidden Markov Chain (HMC) mechanism. The observable variable in this case is going to be the steady state of the system, when the real variable following it is hidden from observation. We assume that the system's steady state changes because of the HMC, but the actual timing of this shift might be slightly slower than in the hidden variable. In other words, from the moment the risk aversion parameter changes, the system will start moving to the new equilibrium. The length of this "switching" period will be dependent on the flexibility of the financial system and on the other parameters involved.

To test whether this theoretical framework is consistent with the empirical data the relationship between the exchange rate and the stock prices will be analyzed through

---

[4] The risk aversion should not be too high since otherwise the behavior of the investors would not converge to any equilibrium. Similarly, the risk aversion is also related inversely with the parameter for price elasticity of forex supply, k. This should also be high enough to allow the system to compensate properly for the changes in stock prices and reach an equilibrium.

a Markov Switching Model with 2 regimes. We will run a regression with the exchange rate returns as a dependent variable and the stock return differentials as the independent one. We expect to find a negative and significant coefficient of the stock returns in both the regimes. If it is so (and we get the same results over the whole sample period), then it is possible to say that the theory by Hau and Rey is consistent with the studied market. Then we will proceed to check whether the series could be the result of two distinct alternating regimes. Our hypothesis is that in the two specifications they will be different and persistent enough to uphold (be consistent with) the assumption of switching risk aversion.

Empirics

Since the theoretical model uses the foreign exchange approach to order flows, the empirical analysis is going to be carried out focusing on the exchange rate series. The object in study will be the exchange rate equation as explained by the difference of the stock returns of the two countries. Formally:

$$\Delta S_t = a + b\left(\Delta R_t^{f*} - \Delta R_t^h\right)/\overline{P}$$

where ΔSt is the first difference of the exchange rate series (in logs) and the regressor is given by the returns on stocks as defined in the model above. According to the above theory the coefficient b should always be significant and negative.

We study the markets of the United Kingdom, Switzerland, pre-Euro Germany and Japan. All exchange rates are against the United States dollar. The sample period is from October 1983 through October 2008. Within this period homogenous data is available for all the countries. For Germany the observations will end on December 1998 to avoid any data misspecification from the introduction of the Euro. The frequency is monthly. The first returns of the stock indexes are used for each country. For the United States the index used is the Standard and Poor's. For United Kingdom it is the FTSE. For Germany we used the DAX and for Switzerland the index is the MSCI. For Japan we used the NIKKEI. All data is taken from Datastream. At first there will be a non-time-varying analysis over the whole sample. This will be needed to verify that the theory used is indeed consistent with the data in the long run. Then it will be interesting to start the time-dependent study. The markets that might be rejecting the theory will also be studied with a time-varying approach. However, the interpretation of the results will be different.

Table 1 shows the results of the regression above. It is possible to see that for three out of the four countries the theoretical framework introduced works perfectly. This is because the table shows the b coefficients to be most always negative and significant (the probability values are shown in parenthesis). The theory predicts indeed a negative correlation between the two variables. The only market that clearly rejects the theory is the Japanese one. The coefficient is not only positive but also insignificant. This means that the negative correlation predicted by the theory is missing.[5]

Since the theory is consistent with the data it is interesting to check whether the time varying characteristic adds to the understanding of the data. In order to test our theory we are going to take a two-step approach. First we are going to check whether it makes any sense to apply a time-dependent specification on these markets. This is done through a Markov Switching Model (2) that is regressed over equation 1. Only if this analysis shows consistent results then it is possible to proceed. Specifically, the regimes will have to be persistent enough to allow for periods of system equilibrium. If this was not to happen it would be hard to see the economic interest of this approach.

[5] This conclusion is based on the data as described. However it matches perfectly with the data shown in Hau, Rey 2006. The correlations between the markets over these samples are -0.26 for the UK, -0.23 for Switzerland, -0.23 for Germany and 0.07 for Japan. This is in line with the coefficient results described above.

Table 1: Coefficients of the regressions on the whole sample

	a	b	R^{2}
UK	-0.051	-0.086	0.057
	(0.48)	(0.00)	
Switzerland	-0.252	-0.086	0.015
	(0.20)	(0.03)	
Germany	-0.002	-0.141	0.051
	(0.32)	(0.00)	
Japan	-0.127	0.049	0.007
	(0.10)	(0.12)	

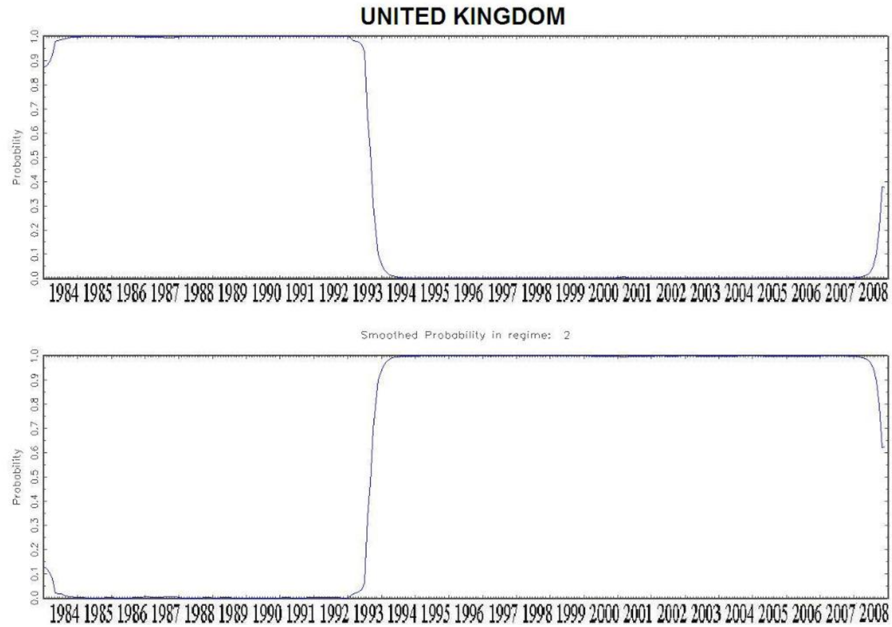


Figure 1: Smoothed probabilities of the two regimes for the British market

If it is shown that the time dependent analysis gives interesting results, step 2 will test whether the structural breaks may be attributed to a change in risk aversion. However (as discussed above) the risk aversion parameter is not observable from these series. Therefore, we will have to devise a way to check if at least the data is consistent with the assumption on risk aversion. This will be done by observing the value of a function of *k* in the periods when the two regimes have reached an equilibrium. This analysis will also show what periods are characterized by an optimistic and by a pessimistic regime.[6]

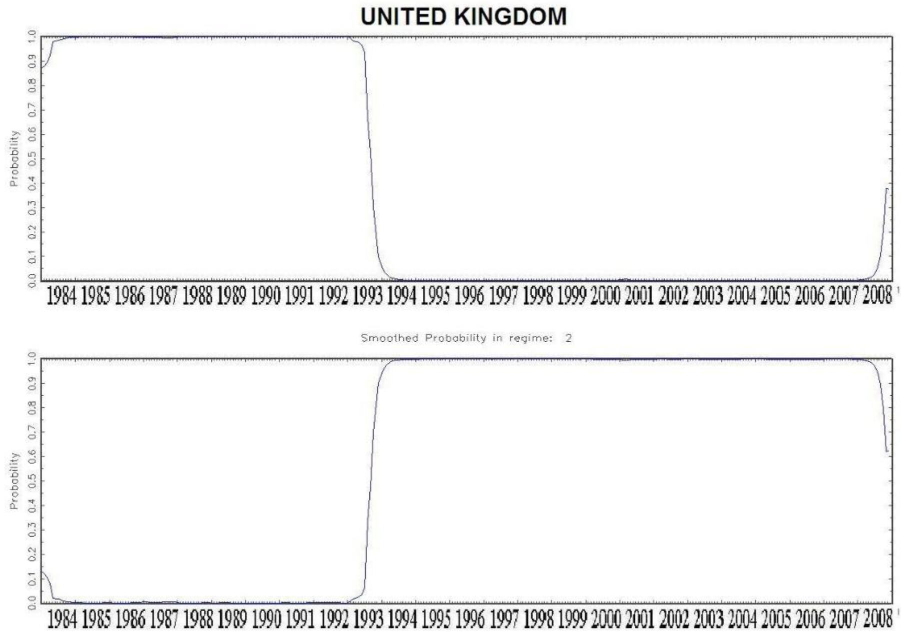


Figure 2: Smoothed probabilities of the two regimes for the British market

The results of the first step are shown in Figures 1 through 4. The figures represent the probability that at every point in time the market is in regime 1 or 2. Obviously the sum of the two is always equal to 1, as the market has to always be in one of the states. When this probability is 1 for the first regime and 0 for the other it is sure that the market is in the first regime. On the other hand, when the probabilities are closer to a 50-50 scenario, it is not possible to tell in what regime the market is in. This could be seen as a period of switching between the two states.

It is interesting to notice how in the markets where our general framework is confirmed, there is a clear distinction between the periods in which the two regimes are acting. A peculiarity emerging from the pictures is also that in all the Hau-Rey type

[6] The robustness checks for this assumptions are carried out and listed in Appendix B.

markets there is one single switch between the regimes. In all cases this switch takes a few time periods to be complete. The length of this "switching" time is due to the characteristics of the economies (i.e. their flexibility in adjusting to a new scenario). It is possible to show two separate time ranges in which the two regimes alternate. This is consistent with the time-varying assumption. Indeed, it is reasonable to believe that if the risk aversion changes in time it is not going to keep switching back and forth. It is going to slowly bring the system towards another equilibrium. This is exactly the picture shown by the empirical analysis.

As shown earlier the Japanese market is not consistent with the general framework used. This means that even in the case that the risk aversion were to change, the system would not necessarily reach a stable equilibrium where the independent variable is relevant. That is because the exchange rate seems not to be dependent on the asset prices as heavily as it is in the other markets (see Table 1 for the overall R^2 values). Then it is natural that also the time-varying study gives different results than for the other countries.

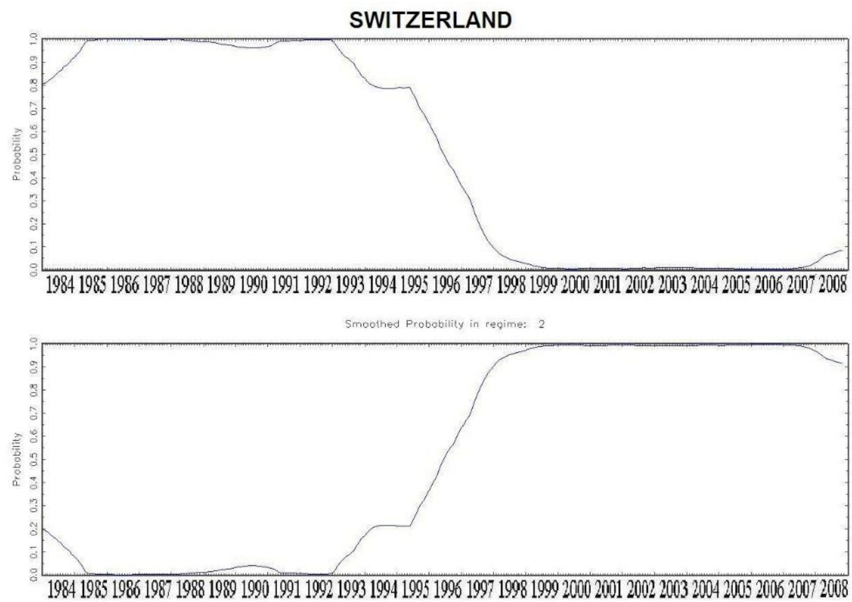


Figure 3: Smoothed probabilities of the two regimes for the Swiss market

Table 2 shows the coefficient *b* in the two regimes of the MSM given that the first regime to chronologically appear in the data is referred to as Regime A and the second as Regime B. It is possible to notice how in every market the two states have very different coefficients. This shows how this analysis adds value in the estimation of the variable.
However, the fact that the markets switch regime does not necessarily prove that the source of the switch is the risk aversion. This is the reason why the second step is needed. We shall look at another statistic shown in Table 2. The variance ratio is formally defined as:

$$Var \text{ Ratio } = \sqrt{\frac{Var(\Delta S_t)}{Var\left(\Delta R_t^{f*}/\overline{P}\right)}}$$

It is a function of the variable *k* described in the previous section. *k* is the elasticity of the exchange rate supply, but in this model is also a function of the risk aversion and of the speed of return to equilibrium. From the table we can see that the Variance Ratio changes considerably between the two regimes, and is always lower in Regime B. This means that the variable *k* becomes higher, which is consistent with a lower risk aversion parameter. In other words the analysis shows that Regime A may be defined a "pessimistic" regime, while Regime B is an "optimistic" regime. As it will be illustrated in the next paragraph, this is coherent with the literature on risk aversion.

The same statistic is also shown for the Japanese market. Of course, in Japan the two periods to measure the ratio were arbitrarily chosen, only to compare with the other markets. The Variance Ratio falls over time also here. This is consistent with the time varying assumption that risk aversion changes normally in economies with capitalistic features. However the effect on the equilibrium of the system may only be seen when the variables considered are important in the determination of the economy steady state.

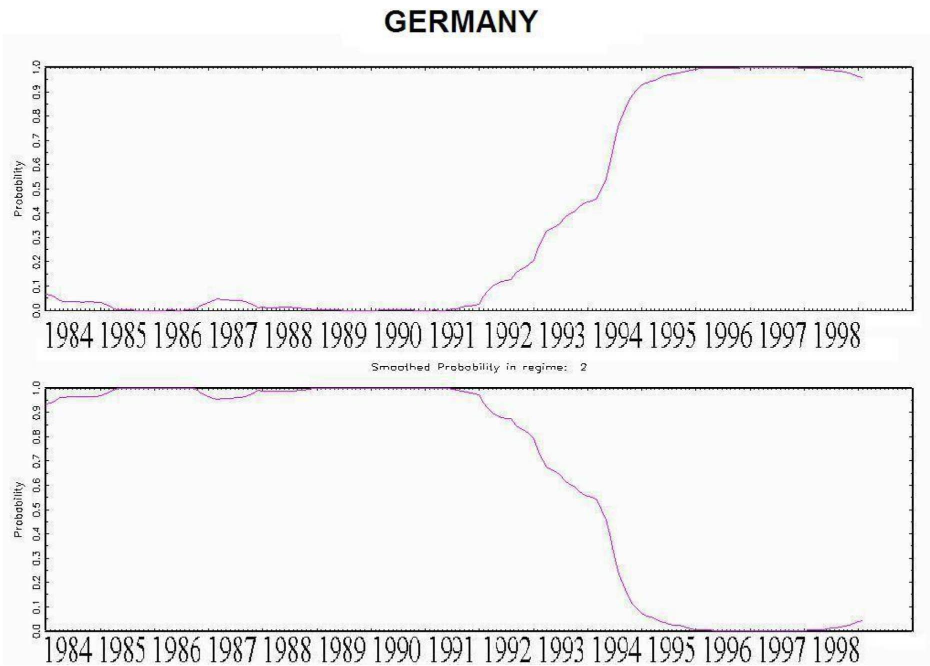


Figure 4: Smoothed probabilities of the two regimes for the German market

For completeness of the study, it is important to also calculate the correlation coefficient in the different regimes. This is also shown in Table 2. The correlation increases in Germany and the UK, while it decreases in Switzerland. The model shows how the theoretical correlation assuming no asymmetries or exogenous shocks would be equal to 1. So the difference in these coefficients over time is attributed to equity market integration Hau and Rey (2006) or to other exogenous factors.

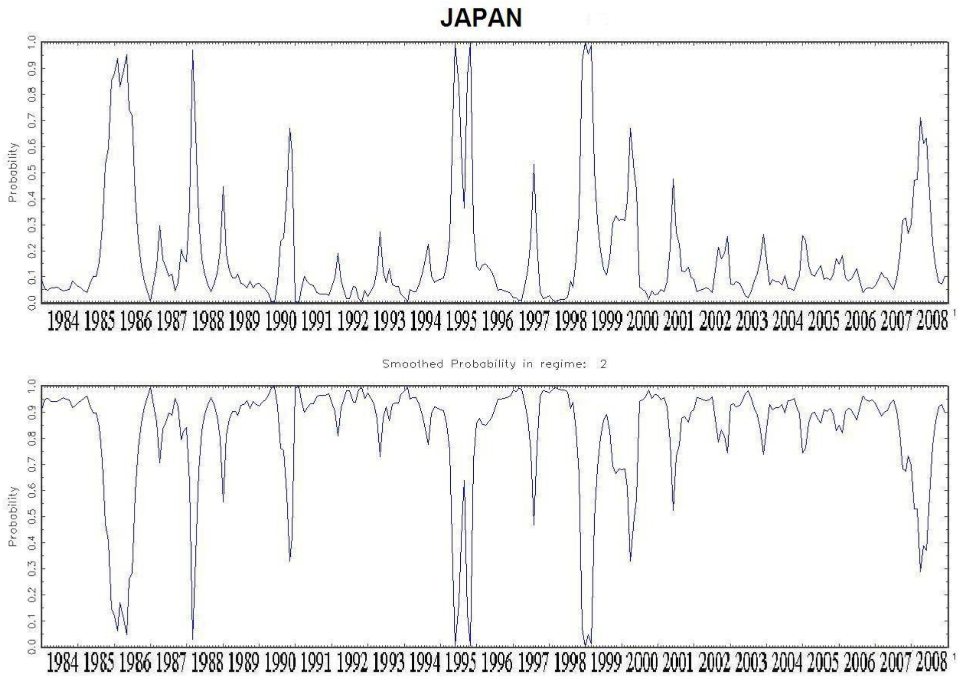


Figure 5: Smoothed probabilities of the two regimes for the Japanese market

Discussion

The theory described in the first part of the paper was tested empirically on the data. The clear result that was obtained in most of the markets is that the data is consistent with the order flow approach and that a time- dependent study upholds the varying risk aversion hypothesis. However, this is not the only new information that came from the data.

First, it may be noticed how in the three countries with a negative correlation, the regimes switch in similar periods. This means that the model has found a generalized lowering of the risk aversion during the nineties. This has begun to change starting in the end of 2007. So, the big question becomes: how does this relate to our literature on the subject?

Risk has always been an interesting topic for research, so the literature on the field is very wide. However, it was not until the 80s and 90s that a varying risk aversion has been studied more in detail Hansen and Singleton (1983); Brunnermeier and Nagel (2004); Brandt and Wang (2002). Specifically, it is now usually accepted that risk aversion may change, although there is not yet agreement on the exact schedule on which it does so. What is more interesting is the fact that the risk aversion literature points to a strong decrease of the perception of risk that started in the nineties (Brandt and Wang, 2002; Brunnermeier and Nagel, 2004).

		Regime A	Regime B
UK	t	Oct 1983 - Nov 1992	Dec 1994 - Nov 2007
	corr	-0.20	-0.40
	coef	-0.21	-0.28
	p-value	0.04	0.00
	var ratio	0.70	0.48
Switzerland	t	Oct 1983 - Oct 1992	Apr 1999 - Nov 2007
	corr	-0.25	-0.19
	coef	-0.14	-0.09
	p-value	0.00	0.01
	var ratio	0.73	0.63
Germany	t	Oct 1983 - Apr 1991	Dec 1995 - Dec 1998
	corr	-0.08	-0.53
	coef	-0.08	-0.33
	p-value	0.17	0.00
	var ratio	0.54	0.45
Japan	t	\\	\\
	corr	\\	\\
	coef	-0.08	0.13
	p-value	0.25	0.00
	var ratio	0.62	0.44

Table 1: Coefficients of the MSM

Furthermore, financial papers and journals have dedicated a lot of attention to varying risk premia in the last years (among the others, see Canto (2005)). It was widely recognized that they have been decreasing sharply starting in the 90s. Risk premia may be related to risk aversion since their fall implies that agents require a lower price for taking the same amount of risk, which could be taken as measure of a decreasing risk aversion.

An important element of the analysis is that it also shows the effects of the current financial crisis. The break originating at the end of 2007 in UK and Switzerland is easy to interpret. After the sub-prime crisis broke out on last year's summer period, people have corrected their risk aversion once again. The system is now switching to the regime that is characterized by a higher risk aversion.

At the light of this research and financial press, it is possible to see the pattern that has been highlighted in this paper. Although the data does not prove the model, it is very coherent with it.

Conclusion

This chapter described an order flow approach to the exchange rate market. This is studied in relationship with the stock prices. The model adopted is of recent development and may be the beginning of a new approach to exchange rate study. The innovation has been to make this approach time-varying.

The analysis that was carried out shows how this modification has potential in explaining the time series. This is due to the clear results provided by the empirical estimation. The regimes found are persistent and in agreement with the theory. Indeed the data shows the changes in the two regimes in a consistent manner with the financial literature and press. That is a crucial result for the theory's consistency.

This shows why the approach taken in this paper is encouraging. However future space for research is left in several aspects. First of all, it would be possible to develop a further specification of the exogenous parameters. Secondly a more explicit process of exchange rate (other than an O-U process) would probably improve the fit of the model. Also, this would render the approach of order flows flexible to be incorporated in exchange rate literature of most schools of thought. Thirdly, characterizing the out of equilibrium dynamics would be interesting to see whether the length of the switching time between regimes is a function of economic variables.

Overall, this theoretical framework seems to be promising and ready for more constraining assumptions. The simplicity of the intuition makes it a powerful tool for the explanation of the exchange rates and the stock prices. The accuracy of the test results testifies to the potential of this theory. The little number of variables used should encourage a wider analysis to include also exogenous observable changes.

References

Abdulnasser Hatemi-Ja and Eduardo Roca. Exchange rates and stock prices interaction during good and bad times: evidence from the asean4 countries. *Applied Financial Economics*, 15:539–546, 2005.

Abdulnasser Hatemi-Ja and Manuchehr Irandoust. On the causality between exchange rates and stock prices: A note. Blackwell Publishers, *Bulletin of economic research*, 54:2, 2002.

Bala Ramasamy and C.H. Matthew Yeung. The causality between stock returns and exchange rates: revisited. Blackwell Publishing Ltd/University of Adelaide and Flinders University, 2005.

Clive W.J. Granger, Bwo-Nung Huang, and Chin Wei Yang. A bivariate causality between stock prices and exchange rates: Evidence from recent asia flu. *Discussion paper*, April 1998.

Martin Evans and Richard Lyons. Order flow and exchange rate dynamics. *Journal of Political Economy*, Vol. 110, 2002. Martin Evans and Richard Lyons. How is macro news transmitted to exchange rates. 2006.

Harald Hau and Helene Rey. Exchange rates, equity prices and capital flows. *review of financial studies*, 2006.

Edoardo Otranto and Giampiero M. Gallo. A nonparametric bayesian approach to detect the number of regimes in markov switching models. *Econometric Reviews*, 21, No. 4:477–496, 2002.

Patricia Alvarez-Plata and Mechthild Schrooten. The Argentinean currency crisis: A markov-switching model estimation. *German Institute for economic research, Berlin, Discussion paper 348*, May 2003

Christopher A. Sims and Tao Zha. Mcmc method for markov mixture simultaneous-equation models: a note. 2004.

Yin-Wong Cheung and Ulf G. Erlandsson. Exchange rates and markov switching dynamics. *CESifo Working paper No. 1348, Cathegory 6: Monetary policy and internatiuonal finance*, December 2004

Zacharias Psaradakis and Nicola Spagnolo. On the determination of markov- switching regimes in markov switching autoregressive models. *Journal of Time series Analysis*, 24, N.2, 2003

Benoit Bellone. Classical estimation of multivariate markov-switching models using msvarlib. July 2005

Hans-Martin Krolzig. Statistical analysis of cointegrated var processes with markovian regime shifts. *Institute of Economics and Statistics and Nuffield College, Oxford, SFB373 Discussion Paper:25*, 1996.

Lars Peter Hansen and Kenneth J. Singleton. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *The Journal of Political Economy*, 1983.

Marcus Brunnermeier and Stefan Nagel. Do fluctuations in wealth generate time-varying risk aversion? micro-evidence on individuals asset allocation. 2004.

Michael W. Brandt and Kevin Q. Wang. Time-varying risk aversion and unexpected inflation. *Journal of Monetary Economics*, 2002

Victor A. Canto. Greenspan the risk manager, 2005. URL http://www.nationalreview.com/nrof_canto/canto200509020857.asp.

Jushan Bai and Pierre Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 1998a.

Jushan Bai and Pierre Perron. Computation and analysis of multiple structural-change models. Center for Interuniversity Research in Quantitative Economics (CIREQ), 1998b.

C Tille and E van Wincoop. International capital flows. 2006.

Michael B Devereux and Alan Sutherland. Solving for country portfolios in open economy macro models. 2007.

Anna Pavlova and Roberto Rigobon. Asset prices and exchange rates. *Review of Financial Studies*, 2007. Robert Kollmann. Explaining international comovements of output and asset returns: The role of money and nomical rigidities. *Journal of Economic Dynamics and Control*, Vol. 25, 2001.

Mark and Nelson. *International Macroeconomics and finance*. Blackwell

Fumio Hayashi. *Econometrics*. Princeton University Press, 2000.

J. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.

C. Engel and J. D. Hamilton. Long swings in the dollar: are they in the data and do the markets know it? *American Economic Review*, 80:689–713, 1990.

J. P. LeSage. *Econometrics toolbox*.

Jesus Munoz and P. Nicholas Snowden. Floating without flotations- the exchange rate and the mexican stock market: 1995-2001. *Journal of International Development*, 18:299–318, 2006.

Soren Johansen. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59, No. 6:1551–1580, November 1991.

K. N. Badhani. Dynamic reloationship among stock prices, exchange rate, and net fii: Investment flow in india. 2005.

Sheng-Yung Yang and Shuh-Chyi Doong. Price and volatility spillovers between stock prices and exchange rates: Empirical evidence from the g-7 countries. *International Journal of Business and Economics*, 3 (2):139–153, 2004.

C.E. Jorion. The exchange rate exposure of us multinationals. *Journal of Business*, 63 (3), 1990.

Torben G. Andersen, Tim Bollerslev, Francis X. Diebold, and Clara Vega. Real-time price discovery in stock, bond, and foreign exchange markets. *National Bureau of economic research, May 2005*.

Carol L. Osler. Currency orders and exchange rate dynamics: An explanation for the predictive success oftechnical analysis. *The Journal of Finance*, October 2003.

Naeem Muhammad and Abdul Rasheed. Stock prices and exchange rates: Are they related? evidence from south asian countries. 2001.

Saadet Kasman. The relationship between exchange rates and stock prices: A causality analysis. 2003.

Desislava Dimitrova. The relationship between exchange rates and stock prices: Studied in a multivariate model. *Issues in Political Economy*, 14, August 2005.

Bhattacharya Basabi and Mukherjee Jaydeep. Causal relationship between stock market and exchange rate, foreign exchange reserves and value of trade balance: A case study for india. 2001.

Issam S.A. Abdalla and Victor Murinde. Exchange rate and stock price interactions in emerging financial markets: evidence on India, Korea, Pakistan and the Philippines. *Applied financial economics*, 7:25–35, 1997.

Brad Comincioli. The stock market as a leading indicator: An application of granger causality. *The University Avenue Undergraduate Journal of Economics*, 1996.

Christopher A. Sims and Tao Zha. Were there regime switches is US monetary policy? 2006.

Changing lenses: micro-based ‘macro’ predictions

Jaap Bos, Ryan van Lamoen and Jiang Yuan

Abstract

This paper examines how firm-level heterogeneity can improve macroeconomic forecasting by integrating micro-based insights into aggregate models. Using quarterly U.S. banking data (1989–2009), we compare a traditional autoregressive model with a dynamic reallocation approach that decomposes industry performance into within-firm, share, covariance, and net-entry effects. Results show that reallocation components—especially within-firm improvements and entry-exit dynamics—enhance predictive accuracy over longer horizons, particularly during periods of economic disruption. The findings highlight dynamic reallocation as a practical tool for translating micro-level dynamics into macro-level forecasts.

> *“The man of system. . . is apt to be very wise in his own conceit; and is often so enamoured with the supposed beauty of his own ideal plan of government, that he cannot suffer the smallest deviation from any part of it. . . He seems to imagine that he can arrange the different members of a great society with as much ease as the hand arranges the different pieces upon a chess-board. He does not consider that in the great chess-board of human society, every single piece has a principle of motion of its own, altogether different from that which the legislature might choose to impress upon it.”*
>
> — Adam Smith in his book The Theory of Moral Sentiments

Introduction

The divide between microeconomics and macroeconomics, although a relatively modern development, has become a defining characteristic of contemporary economic thought. Until the early twentieth century, economics was largely a unified field, with theorists such as Adam Smith and David Ricardo addressing questions that we would now classify under both domains (Rodrigo, 2017). The formal separation that emerged following the Great Depression, solidified by Keynesian and later neoclassical models, created distinct theoretical and methodological tracks. While microeconomics coalesced around a unified framework grounded in utility maximization, firm behavior, and market equilibrium, macroeconomics evolved into a fragmented field characterized by competing schools of thought, such as New Keynesian and New Classical approaches.

This divergence has had profound implications for economic research and policymaking. Microeconomics has generally maintained a stable core of principles and empirical methodologies. In contrast, macroeconomics has been shaped by shifting paradigms, recurrent crises of confidence, and persistent methodological disputes. As Storm (2021) argues, the dominance of dynamic stochastic general equilibrium (DSGE) models in mainstream macroeconomics has limited the discipline's ability to integrate richer, more empirically grounded insights—especially those emerging from microeconomic analysis.

Over the past two decades, however, a growing body of work has begun to challenge this disciplinary bifurcation. New empirical strategies, the increased availability of granular micro-level data, and advances in computing power have enabled researchers to explore macroeconomic questions through a microeconomic lens. The 2025 International Monetary Fund (IMF) conference, *From Micro to Macro: Leveraging Microdata to Address Macroeconomic Issues*, exemplifies this shift in emphasis. The event underscores how administrative records, firm- and household-level surveys, and expectations data can be used to better inform macroeconomic theory and policy (International Monetary Fund, 2025).

Despite this progress, the integration of microeconomic insights into macroeconomic modeling remains uneven. Blanchard (2018) argues that while macroeconomic analysis must be grounded in general equilibrium thinking, not all models require strict adherence to micro foundations. He contends that partial-equilibrium analysis, often based on micro-level evidence, is crucial for understanding specific mechanisms. However, the leap from partial to general equilibrium effects requires caution and a well-structured synthesis—something that has proven elusive in mainstream macroeconomic practice.

A central point of divergence between micro and macro lies in their empirical aims—particularly in prediction. Microeconomists typically aim to estimate causal effects within well-defined, often tightly scoped contexts, using natural experiments, field studies, or randomized controlled trials. Predictive accuracy in microeconomics is generally localized and policy-specific: What happens if a minimum wage is introduced, a tax is changed, or a firm faces new competition? These studies are characterized by strong identification strategies, heterogeneity in behavior, and a deep concern for institutional context.

Macroeconomics, by contrast, has historically pursued broader, system-wide predictions—such as GDP growth, inflation dynamics, or the effect of monetary policy shocks—often using aggregate data and structural models. However, the forecasting performance of macroeconomic models has been widely criticized, particularly after the global financial crisis. DSGE models, despite their theoretical rigor, have underperformed in out-of-sample prediction and have struggled to account for key empirical regularities without ad hoc modifications (Storm, 2021). While purely statistical approaches such as vector autoregressions (VARs) offer improved predictive power, they often lack the interpretability and policy counterfactuals that structural models promise.

This gap in predictive performance raises questions about the empirical foundations of macroeconomic modeling. The increasing availability of rich microdata—such as administrative tax records, credit registries, and firm-level panel data—offers a unique opportunity to recalibrate macroeconomic models using evidence from the ground up. As the IMF conference notes, these micro sources are no longer peripheral to economic research but are central to answering core macroeconomic questions. Models that ignore the behavioral heterogeneity revealed by microdata risk misrepresenting the aggregate dynamics they aim to explain.

Nonetheless, simply adding more microdata to existing DSGE or VAR frameworks is unlikely to deliver the hoped-for forecasting gains. What is needed is a way to aggregate heterogeneity that is both empirically measurable and theoretically interpretable. In this paper, we make the case that *dynamic reallocation*—the continuous reshuffling of resources across heterogeneous firms—provides precisely this missing layer, translating firm-level heterogeneity into a set of macro-relevant statistics that can be observed at high frequency.

Dynamic Reallocation as the Missing Link

Dynamic reallocation models, first formalized by Baily, Hulten, and Campbell (1992) and subsequently refined by Haltiwanger (1997), Stiroh (2000), and Decker et al. (2016), decompose aggregate outcomes into four conceptually distinct margins: within-firm improvements, between-firm covariance, market-share shifts, and net entry or exit. Taken together, these margins trace how resources are continuously reassigned across producers in response to shocks, policy changes, and technological innovations.

Why is this decomposition crucial for macro prediction? In standard representative-agent frameworks, aggregate variables inherit the dynamics of a stylized average firm or household. By construction, such models overlook the fact that recessions, recoveries, and structural transformations are, to a large extent, reallocation events. Productivity gains, for example, arise not only because existing firms innovate (a within effect) but also because less-productive firms lose market share or exit entirely (the share and entry effects). Ignoring these margins can therefore lead to systematic forecast errors, especially around turning points when exit, entry, and market-share churn accelerate.

Dynamic reallocation models offer two forecasting advantages. First, they translate micro-level heterogeneity into a parsimonious set of aggregate statistics that can be measured with administrative data at quarterly or even higher frequencies. Second, the decomposition is structurally interpretable: It allows policymakers to map prospective interventions—such as credit guarantees for small firms or reforms to bankruptcy laws—onto specific reallocation channels and, by extension, onto macro aggregates.

Empirical Strategy and Preview of Results

We operationalize these ideas by building a panel of U.S. commercial banks spanning 1984–2004 and estimating a dynamic reallocation model at the quarterly frequency. The banking sector provides an ideal testing ground: The entry and exit of banks, mergers, and shifts in loan portfolios generate pronounced reallocation dynamics, yet rich supervisory data allow us to observe firm-level balance sheets and performance metrics. We then embed the resulting decomposition in a one-step-ahead forecasting equation that links lagged reallocation components to future industry-level productivity growth.[1]

Two findings stand out. First, historical reallocation patterns—particularly the covariance and entry margins—possess substantial predictive power above and beyond aggregate controls. Second, forecasts that exploit the full decomposition outperform those based on aggregate autoregressions in both in-sample cross-validation and genuine out-of-sample tests (1999Q1–2004Q2). The gains are especially pronounced around the 1990–1991 recession and the 2001 dot-com bust, periods in which entry and exit dynamics were unusually volatile.

These results confirm that micro-based reallocation statistics are not merely descriptive; they are causally and predictively informative about future macro-

[1] Hereafter, we use industry growth, industry growth in performance and industry growth in productivity interchangeably.

outcomes. More broadly, our framework illustrates how dynamic reallocation can serve as an intermediate layer in the construction of empirically disciplined macro models: rich enough to respect heterogeneity yet compact enough for tractable forecasting and counterfactual analysis.

Road Map of the Paper

The next section develops the theoretical underpinnings of the dynamic reallocation decomposition and situates it within the broader literature on endogenous growth and misallocation. The section that follows summarizes key stylized facts from our bank-level panel and motivates the forecasting specification. The subsequent section presents the empirical results, including robustness checks and a comparison with benchmark VAR and DSGE forecasts. The section after that discusses policy implications and the final section concludes by outlining avenues for extending the approach to other sectors and macro variables.

Data and Methodology

Data

The dataset contains quarterly bank data at the firm-level from the first quarter of 1989 up and until the first quarter of 2009. Data are gathered from the Call Reports of Income and Condition provided by the Federal Reserve System. The Call Reports include complete balance sheet and income statement data for each bank. All data are expressed in 1989 U.S. dollars. The bank-level data is aggregated to obtain time series data at the industry-level

Measuring performance: Return on Assets

Two often used profitability measures in banking are the return on equity (ROE) and the return on assets (ROA). However, the ROA is often preferred for comparisons by size groups, since larger banks tend to be more highly leveraged than smaller banks (Boyd and Graham, 1991). Although a group size comparison is not possible with aggregated industry-level data, the ROA is used to measure profitability in this paper, because the size of banks is increasing over time and the number of banks is declining in the ongo-ing consolidation trend. Hence, the banks in the industry may become more leveraged over time and therefore distort comparisons across time periods if the ROE is used as a performance measure at the industry-level. Industry ROA at time t is defined as the net income after taxes divided by the total assets of the industry:

$$R_t = \frac{I_t}{A_t} = \frac{\sum_{i=1}^{n_t} I_{it}}{\sum_{i=1}^{n_t} A_{it}},$$

where It and At are the net income after taxes and the total assets at the industry-level, respectively. Both It and At are obtained by aggregating these variables over all banks that operate in each period t.

Table: Descriptive statistics

VARIABLE	OBS.	MEAN	STD. DEV.	MINIMUM	MAXIMUM
RETURN ON ASSETS	81	0.0065595	0.0034663	0.0009258	0.0132367
ΔRETURN ON ASSETS	80	-0.0000165	0.0044902	-0.0096334	0.0036015

The above table shows the descriptive statistics of the ROA. The average ROA and average change in the ROA are 0.0065595 and -0.0000165, respectively.

The Autoregressive model

The main purpose of this paper is to examine the performance of two models to predict changes in the ROA.[2] In the first prediction model, we use an autoregressive process of order p (AR(p)) to predict changes in the aggregate performance of the U.S. banking industry:

$$\Delta R_t = \delta + \theta_1 \Delta R_{t-1} + \cdots + \theta_p \Delta R_{t-p} + \varepsilon_{1t},$$

where Rt represents the return on assets, δ a constant term and ε_{1t} is assumed to be a white noise process.[3] In this AR(p) process, the change in the ROA in the current period is a function of changes in the ROA in the past. The Akaike Information Criterion

[2]Hence, we are not interested in predicting levels of the ROA.

[3] Depending on the influence of seasonality in the series, the equation can be extended by including seasonal dummies.

(AIC) and the Bayesian Information Criterion (BIC) are used to determine the lag order of the AR(p) process.[4]

The Dynamic Reallocation Model

In the second prediction model, we use past values of the decomposition components in the Dynamic Reallocation Model to predict industry performance. The Dynamic Reallocation Model used by Stiroh (2000) begins with the measurement industry performance measurement (profitability). In this paper, we use the ROA and denote the profitability in the banking industry in time t as Rt. Therefore, ΔRt = Rt −Rt−1 indicates the growth of the industry's profitability between time t and t-1. Industry performance R_t can be rewritten as:

$$R_t = \sum_{i=1}^{n_t} \frac{I_{it}}{A_{it}} \frac{A_{it}}{\sum_{i=1}^{nt} A_{it}} = \sum_{i=1}^{n_t} R_{it}\theta_{it}$$

The above equation shows that the industry ROA can be rewritten as a weighted average of the ROA of the individual banks, where the market share weights θ are based on total assets. Then, we can decompose the changes in industry performance (ΔRt) into the following components:

$$\Delta R_t = \sum R_{it}\theta_{it-1} - \sum R_{it-1}\theta_{it-1} = \sum_{i\in S_t}[\Delta R_{it}\theta_{it-1}] + \sum_{i\in S_t}[\Delta R_{it}(R_{it} - R_{t-1})] + \sum_{i\in S_t}[\Delta R_{it}\Delta\theta_{it}] + \sum_{i\in E_t}[\theta_{it}(R_{it} - R_{t-1})] - \sum_{i\in X_{t-1}}[\theta_{it-1}(R_{it-1} - R_{t-1})]$$

The within effect ($\sum_{i\in S_t}[\Delta R_{it}\theta_{it-1}]$) in this equation measures the contribution of the surviving banks between two subsequent periods (subset S from the sum of all banks) to changes in the ROA at the industry level. The share effect ($\sum_{i\in S_t}[\Delta R_{it}(R_{it} - R_{t-1})]$) represents the impact of changes in the relative bank size on performance growth by surviving banks. The covariance effect ($\sum_{i\in S_t}[\Delta R_{it}\Delta\theta_{it}]$) contains the contribution to changes in the industry ROA by surviving banks due to changes in performance and changes in the relative size of banks. The entry and exit effect are given by

[4] The AIC and BIC provide a trade-off between goodness-of-fit and the parsimony of the model. Models with the lowest AIC and BIC are usually preferred. The penalty for an increasing number of regressors is larger in the BIC.

$\sum_{i\in E_t}[\theta_{it}(R_{it} - R_{t-1})]$ and $\sum_{i\in X_{t-1}}[\theta_{it-1}(R_{it-1} - R_{t-1})]$, respectively. The entry effect captures the contribution of entering banks (subset E from the sum of all banks) and the exit effect the contribu-tion of exiting banks (subset X from the sum of all banks) to the growth in the industry-level performance. The net-entry effect is obtained by subtracting the exit effect from the entry effect. The decomposition components reveal more information on the sources of the growth in industry performance.

By denoting the within effect, share effect, covariance effect and net entry effect as *W*, *SH*, *C* and *NE* respectively, we specify the following model to predict changes in industry performance:

$$\Delta R_t = \alpha + \gamma_1 W_{t-1} + \gamma_2 SH_{t-1} + \gamma_3 C_{t-1} + \gamma_4 NE_{t-1} + \cdots + \gamma_k W_{t-p} + \gamma_{k+1} SH_{t-p} + \gamma_{k+2} C_{t-p} + \gamma_{k+3} NE_{t-p} + \varepsilon_{2t},$$

where α represents a constant term and ε2t is assumed to be a white noise process.5 The AIC and BIC are used to determine the lag structure of the decomposition components. In this equation, changes in the ROA are a function of past information of the within effect, share effect, covariance effect and net entry effect. Each of these decomposition components may provide information on the sources that drive performance growth in future periods and are therefore treated as separate stochastic processes. Hence, we allow for differences in the parameters of the decomposition components and consequently different parameter estimates between the models in the first and last equation, above. However, the same parameter estimates from the models can be obtained by imposing additional restrictions on the parameters. Specifically, the conditions that γ1 = γ2 = γ3 = γ4 for the decomposition components in period t-1 must hold up and including $\gamma_k = \gamma_{k+1} = \gamma_{k+2} = \gamma_{k+3}$ for the decomposition components in period t-p. That is, the parameters of the decomposition components in the past may differ between periods, but not within a certain period. Basically, both models are the same if the decomposition components are substituted in the AR(p) model in the second equation. However, we apply the OLS

[5] Depending on the influence of seasonality in the series, this equation can also be extended by including seasonal dummies.

estimator on both equations (i.e., without imposing parameter constraints on the last equation) and compare the performance of these prediction models.

Examining out of sample performance

The main purpose of the paper is to compare the out of sample prediction performance between the models in both equations. The preferred prediction model minimizes the expected loss:

$$\min_{\hat{R}} E\left[L(e) \mid x\right]$$

where $L(e)=L(R-\hat{R})$. An often-used loss function is the mean-square loss $(L(e) = \overline{e^2})$. In this paper, we also use the Mean Squared Error (MSE) to compare the prediction performance of the two models. This performance measure penalizes larger deviations of the predictions from the actual values more than smaller deviations of the predictions from the actual values. We follow two procedures to compare the two prediction models. First, the performance of the two models are compared by utilizing the MSE and different out of sample periods. The out of sample windows used in this paper are 1 quarter, 4 quarters, 8 quarters and 12 quarters. Furthermore, we also predict 1 quarter ahead based on a moving information window of 40 quarters. Second, we examine the prediction performance based on the MSE, a fixed out of sample period of 1,4,8 and 12 quarters and various information windows. In this approach, the initial information window is based on the whole series. Subsequently, the information window is reduced up and until a minimum of 30 observations.

Results

This section provides the results to compare the prediction performance of the AR(p) model in the second equation and last equation based on lags of decomposition components. First, we present the estimation results based on the whole sample period. Second, we compare the Mean Squared Error based on several specifications with various out of sample and information windows.

Stationarity, slope homogeneity and in sample prediction performance

Augmented Dickey Fuller (ADF) tests are used to examine the stationarity properties of the series according to the Pantula principle, where the ADF is applied first to the series differenced additional times. Different lag orders are included to examine the sensitivity of the results.[6] The null hypothesis that ΔROA contains a unit root is rejected at the 5% in the ADF tests including up to four lags (in second differences).[7] The null hypothesis that the decomposition components contain a unit root is also rejected at the 5% level in most ADF tests.[8] Hence, we assume that all the series follow a stationary process.

Based on the AIC and BIC, four lags are included in both prediction models.[9] The next table gives an overview of the OLS results, including a constant and quarter dummies. The results show that in the AR(4) model, only the first lag and fourth lag are significant at the 1%level. The third lag is only significant at the 10% level. The results of the model based on lags of the decomposition components show that most of the share, covariance and net entry effect are insignificant. The first lag and fourth lag of the within effect are significant at the 1% level, and the second and third lag significant at the 5% and 10%, respectively. The fourth lag of the net entry effect is only significant at the 10% level. These results indicate that the within effect is mostly explaining future changes in ROA. This means that surviving banks in two consecutive periods are mostly contributing to the growth in industry performance in subsequent periods.

Table: In sample performance

MSE	OBSERVATIONS	MEAN	STD. DEV.	MINIMUM	MAXIMUM
ORIGINAL MODEL	76	8.95e-07	2.32e-06	1.73e-14	0.000017
REALLOCATION MODEL*	76	7.74e-07	1.93e-06	3.10e-11	0.0000137

The asterisk * indicates the preferred prediction model.

---

[6] We also performed tests with and without a constant term.
[7] This result was obtained in the tests including a constant term.
[8] This result does not hold for more than five lags for the within effect, seven for the share effect and six for the covariance effect.
[9] However, the AIC indicates an optimal lag length of five lags in equation (5a). We use four lags in this model, since the BIC indicates four lags as optimal and to keep the models as comparable as possible.

We also performed an F-test to examine whether the coefficients of the decomposition components differ significantly from each other within a certain period. Thus we perform the test for the lags in t-1 until t-p separately.[10] The F-tests indicate that the coefficients of the decomposition components within each periods are not significantly different.[11] Using Constraint Least Squares, where the coefficients of the decomposition components are constrained to be equal within a certain period yields exactly the same parameter estimates as performing OLS on the second equation.

[10] Hence, the test is not performed jointly on decomposition components that differ in lag order.
[11] The F-statistics for the lags in period t-1, t-2, t-3 and t-4 are 0.67, 0.83, 0.74 and 1.16, respectively.

Table: Estimation results based on all time periods

DEPENDENT VARIABLE:	EQUATION (1)	EQUATION (2)
ΔROA_{t-1}	***-0.2846613	
	(0.1007537)	
ΔROA_{t-2}	-0.1660211	
	(0.1065716)	
ΔROA_{t-3}	*-0.1825072	
	(0.1029757)	
ΔROA_{t-4}	***0.569181	
	(0.1025131)	
$\Delta WITHIN_{t-1}$		***-0.295788
		(0.1093018)
$\Delta WITHIN_{t-2}$		**-0.2452517
		(0.1189254)
$\Delta WITHIN_{t-3}$		*-0.2308011
		(0.1177926)
$\Delta WITHIN_{t-4}$		***0.5141164
		(0.1198773)
$\Delta SHARE_{t-1}$		-2.103799
		(1.721543)
$\Delta SHARE_{t-2}$		-0.9568793
		(1.835702)
$\Delta SHARE_{t-3}$		1.193453
		(1.729471)
$\Delta SHARE_{t-4}$		-0.33647
		(1.674541)
$\Delta COVARIANCE_{t-1}$		-0.97023
		(1.010806)
$\Delta COVARIANCE_{t-2}$		-0.2550844
		(1.089773)
$\Delta COVARIANCE_{t-3}$		0.4832815
		(1.075059)
$\Delta COVARIANCE_{t-4}$		0.1323154
		(1.023386)
$\Delta ENTRY_{t-1} - EXIT_{t-1}$		-0.9824688
		(1.17426)
$\Delta ENTRY_{t-1} - EXIT_{t-2}$		0.8218951
		(1.169362)
$\Delta ENTRY_{t-1} - EXIT_{t-3}$		2.040914
		(1.971076)
$\Delta ENTRY_{t-1} - EXIT_{t-4}$		*3.586414
		(2.054067)
CONSTANT	**-0.0016854	**-0.0018436
	(0.0007005)	(0.0007896)
OBSERVATIONS	76	76
SEASONAL EFFECTS	yes	yes

Standard errors (in parentheses) and significance levels: * — 10% level, ** — 5% level, *** — 1% level. No evidence of heteroskedasticity or serial correlation (Breusch-Pagan and Breusch-Godfrey tests, respectively).

The next table shows the MSE based on the OLS estimations of both equations for the whole sample period. These results indicate that the model based on lags of the decomposition components performs better (based on the MSE criterion).

Table: Out of sample performance

MODEL	Q	MEAN	SD	MINIMUM	MAXIMUM
ORIGINAL MODEL*	1	0.0000378		0.0000378	0.0000378
REALLOCATION	1	0.000044		0.000044	0.000044
ORIGINAL *	4	9.86e-06	7.04e-06	4.44e-06	0.0000201
REALLOCATION	4	0.0000104	5.30e-06	5.98e-06	0.0000181
ORIGINAL	8	6.20e-06	5.96e-06	9.60e-08	0.0000179
REALLOCATION *	8	5.86e-06	5.40e-06	1.00e-08	0.0000149
ORIGINAL	12	4.15e-06	5.55e-06	5.06e-08	0.0000182
REALLOCATION *	12	3.96e-06	5.08e-06	3.16e-10	0.0000154

The asterisk * indicates the preferred prediction model.

Prediction performance: various out of sample periods

This section provides the results based on changing the out of sample period. We examine the out of sample performance based on an out of sample period of 1 quarter, 4, 8 and 12 quarters. The table above provides an overview of the results. The AR(4) model performs better in terms of the MSE with an out of sample window of 1 and 4 quarters, while the model based on the decomposition components performs better with an out of sample window of 8 and 12 quarters. The results based on the MSE provides some evidence that the AR(4) model provides better short-run predictions, while the model based on lags of the decomposition components provides better predictions over longer periods.

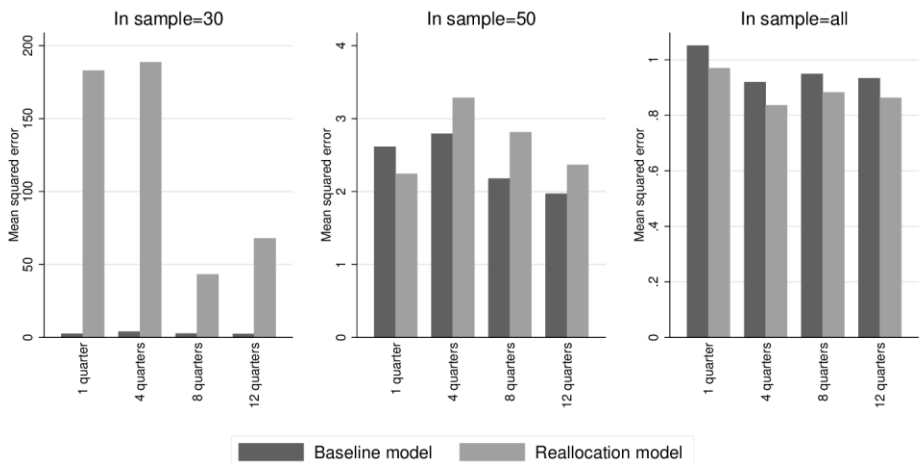
We also performed regressions based on 40 quarters with a moving window of 1 quarter to predict one quarter ahead. The MSE for the AR(4) model in equation (2) and the model in equation (5a) based on the 36 out of sample squared prediction errors are 0.00000323 and 0.0000107, respectively. This implies that the AR(4) model outperforms the model based on lags of the decomposition components if one is interested in predicting the performance growth one quarter ahead.

Prediction performance: various information windows

In this section, we compare the prediction performance of both models by reducing the information window. The out of sample performance is analyzed for an out of sample period of 1 quarter, 4 quarters, 8 quarters and 12 quarters. Three estimations are per-formed for each model and out of sample period. The three estimations are based on the full (available) series, a minimum of 30 observations and a midpoint amount of observations.

The table below provides an overview of the results. The model based on the decomposition components performs better in terms of the Mean Squared Error when using the full available series and 1 quarter as an out of sample period. However, using 30 observations leads to a better performance of the AR(4) model. This finding is similar when using 4 quarters, 8 quarters and 12 quarters as an out of sample period. The model based on the decomposition components seems to perform better when larger information windows are available. The AR(4) model performs better when shorter series are available. An explanation for this result is that the AR(4) model contains fewer explanatory variables.

Figure: Out of sample performance



The AR(4) model has more degrees of freedom in the estimations compared to the model based on decomposition components. This may play an important role for the estimates when short time series are available.

Table: Out of Sample Performance (Changing Information Window)

MODEL	Q	OBS.	MEAN	SD	MINIMUM	MAXIMUM
ORIGINAL	1	76	0.000037			
REALLOCATION *	1	76	0.000037			
ORIGINAL	1	53	0.000071			
REALLOCATION *	1	53	0.000068			
ORIGINAL *	1	30	0.000071			
REALLOCATION	1	30	0.000088			
ORIGINAL	4 s	73	0.000010	0.000007	0.000004	0.000020
REALLOCATION *	4 s	73	0.000009	0.000006	0.000005	0.000018
ORIGINAL *	4 s	52	0.000015	0.000014	0.000004	0.000035
REALLOCATION	4 s	52	0.000016	0.000014	0.000004	0.000035
ORIGINAL *	4 s	30	0.000018	0.000018	0.000004	0.000043
REALLOCATION	4 s	30	0.000033	0.000037	0.000001	0.000086
ORIGINAL	8 s	69	0.000006	0.000006	0.000000	0.000017
REALLOCATION *	8 s	69	0.000006	0.000006	0.000000	0.000018
ORIGINAL *	8 s	50	0.000008	0.000008	0.000000	0.000021
REALLOCATION	8 s	50	0.000009	0.000009	0.000000	0.000024
ORIGINAL *	8 s	30	0.000009	0.000010	0.000000	0.000024
REALLOCATION	8 s	30	0.000011	0.000012	0.000000	0.000031
ORIGINAL	12 s	65	0.000004	0.000005	0.000000	0.000018
REALLOCATION *	12 s	65	0.000004	0.000005	0.000000	0.000018
ORIGINAL *	12 s	48	0.000005	0.000007	0.000000	0.000020
REALLOCATION	12 s	48	0.000005	0.000007	0.000000	0.000020
ORIGINAL *	12 s	30	0.000006	0.000008	0.000000	0.000022
REALLOCATION	12 s	30	0.000007	0.000010	0.000000	0.000034

The asterisk * indicates the preferred prediction model.

Conclusion

This paper demonstrates that incorporating firm-level dynamics into macroeconomic forecasting can yield tangible improvements in predictive performance, particularly over medium- to long-term horizons. By decomposing industry performance into interpretable reallocation components—within-firm improvements, market share shifts, covariance effects, and net entry—we show that micro-level heterogeneity contains valuable signals about aggregate outcomes that standard time-series models often miss.

For macroeconomists, the key takeaway is that predictive models need not choose between structural interpretability and empirical accuracy. Dynamic reallocation provides a middle ground: it captures the granular adjustments that underpin aggregate change, while remaining parsimonious enough for forecasting and policy applications. Our results suggest that periods of economic transition—such as recessions or sectoral shifts—are especially dependent on the composition and evolution of micro-level actors.

Rather than viewing microdata as a supplement to traditional macro tools, we argue that it should play a foundational role. Reallocation-based statistics offer a scalable way to track real-time adjustments in the economy, inform countercyclical policy, and refine the assumptions embedded in structural models. Future work should explore how these insights generalize across sectors and can be embedded more directly into general equilibrium frameworks.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5), 1977–2016.

Bachmann, R., & Bayer, C. (2013). Uncertainty business cycles: Evidence from industry-level data. *Econometrica*, 81(3), 765–789.

Baily, M. N., Hulten, C., & Campbell, D. (1992). Productivity dynamics in manufacturing establishments. *Brookings Papers on Economic Activity: Microeconomics*, 1992(1), 187–267.

Blanchard, O. (2018). On the future of macroeconomic models. *Oxford Review of Economic Policy*, 34(1–2), 43–54.

Blundell, R., Pistaferri, L., & Preston, I. (2009). *Microeconometric models of consumer behavior and macroeconomic applications*. Cambridge University Press.

Boulakia, J. D. C. (1971). Ibn Khaldun: A fourteenth-century economist. *The Journal of Political Economy*, 79(5), 1105–1118.

Choi, H., & Varian, H. R. (2012). Predicting the present with Google Trends. *Economic Record*, 88(1), 2–9.

Decker, R. A., Haltiwanger, J., Jarmin, R., & Miranda, J. (2016). Declining business dynamism: Implications for productivity? *Brookings Papers on Economic Activity*, 2016(1), 1–38.

Diebold, F. X. (1998). The past, present, and future of macroeconomic forecasting. *Journal of Economic Perspectives*, 12(2), 175–192.

Durlauf, S. N., & Quah, D. T. (1998). The new empirics of economic growth. *NBER Working Papers*, No. 6422.

Easterly, W. (2002). *The elusive quest for growth: Economists' adventures and misadventures in the tropics*. MIT Press.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3), 733–772.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3), 535–574.

Haltiwanger, J. (1997). Measuring and analyzing aggregate fluctuations: The importance of building from microeconomic evidence. *Federal Reserve Bank of St. Louis Review*, 79(3), 55–78.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, 95(2), 347–361.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long-run equilibrium. *Econometrica*, 60(5), 1127–1150.

International Monetary Fund. (2025, April). *From micro to macro: Leveraging microdata to address macroeconomic issues* [Conference]. Washington, DC, United States.

Jeon, Y., & Miller, S. M. (2005). An 'ideal' decomposition of industry dynamics: An application to the nationwide and state level U.S. banking industry. *University of Connecticut, Department of Economics Working Papers*, 2005-25.

Jordà, Ò., Schularick, M., & Taylor, A. M. (2019). The real effects of credit supply shocks. *Review of Economic Studies*, 86(1), 312–352.

Mian, A., & Sufi, A. (2010). Household leverage and the recession of 2007 to 2009. *IMF Economic Review*, 58(1), 74–117.

Rodrigo, C. (2017). Micro and macro: The economic divide. *Finance and Development*, 54(2).

Rodrigo, C. (2017). The evolution of micro-macroeconomic separation in economic thought. *History of Economic Ideas*, 25(2), 29–50.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1), 65–94.

Stiroh, K. J. (2000). Compositional dynamics and the performance of the U.S. banking industry. *Federal Reserve Bank of New York Staff Reports*, 98, 25.

Stiroh, K. J. (2000). How did the productivity resurgence happen? *Current Issues in Economics and Finance*, 6(8), 1–6.

Storm, S. (2021). Cordon of conformity: Why DSGE models are not the future of macroeconomics. *International Journal of Political Economy*, 50(2), 77–98.

Storm, S. (2021). The variety of capitalism approach to economic policy. *Journal of Economic Surveys*, 35(1), 155–180.