

# Lost in deflation? Unit labour costs and the (mis)measurement of price competitiveness

Michał Gądomski

Université Rennes 2 – LiRIS

## Introduction

Taking as a point of departure the literature on diverging economic trajectories in Europe at the intersection of comparative political economy and international economics, this paper aims to investigate some weaknesses of the price competitiveness indicators – the deflators of Real Effective Exchange Rates, REERs – used in this literature. After a literature review outlining the arguments, data and methods mobilised in this debate, we first define what we call the fundamental problem in the measure of price competitiveness: economies trade different goods, different qualities of similar goods, different tasks within GVCs, which makes direct comparisons of relative prices impossible. We then outline two ways out of the difficulty thus framed: comparing price changes and comparing a single key price, that of labour.

The first track is that implicitly taken when using price indices as deflators, assuming the broader price movements they measure are a good proxy for the movement of the relevant relative prices. As we outline in a third section, there are a priori reasons to doubt the costs faced by exporting firms closely track general or sectoral price indices.

The second track is that taken when using Unit Labour Costs (ULCs) as deflators. Despite the theoretical appeal of this approach, it poses its own set of difficulties, which we outline in section four. First, ULCs only represent a fraction of firms' cost structures. Second, they are potentially endogenous to export performance – they incorporate value added, which is procyclical and affected by exports. Given that wages and employment tend to lag value added, a better-performing sector should mechanically produce a reduction in ULCs. Finally, they are in fact only in part a measure of wage costs: they are a wage share multiplied by a price index, which causes their movements to be dominated, over longer periods, by the price index component as wage shares tend to be relatively stable. ULC-based Real Effective Exchange Rates (ULC-REERs) incorporate a further source of variance, the nominal exchange rate, which together with prices dominates their movements.

The significance of these issues for the estimation of export equations is then analysed using Monte Carlo simulations with fictitious data and running a range of econometric methods used in the literature: OLS on differences, ARDL models with bounds tests, panel fixed effects and mean group panel ARDL. We construct a simple data generating process (DGP) with export equations that are functions of foreign demand and exogenous shocks, employment that adjusts to output with a lag and wages which are a random walk with a trend equalling that of productivity.

Our analysis shows that the endogeneity of ULCs to value added can create a statistically significant relationship between export growth and ULCs – despite the complete absence, by construction, of any wage effects on exports – when using OLS on differences, fixed effects models and some specifications of mean group panel ARDL, but not single-unit ARDL models with bounds tests. The ARDL, though also biased, remains robust in terms of false positives as the underlying labour-hoarding mechanism by construction only affects ULCs transitorily. Introducing into the DGP a Verdoorn effect, with export success durably increasing productivity and thus lowering the wage share, barely changes this result. Symmetrically, when a wage-share elasticity of realistic size (proportional to labour’s share in export costs) is introduced into the DGP, the ARDL rarely detects it through the ULC-REER, as the wage share accounts for only about a fifth of the ULC-REER’s variance. This leads us to conclude that the elasticities detected in the literature are likely driven by the price components of the ULC-REER rather than wage or productivity dynamics. The paper concludes with a discussion of the implications of these results for the analysis of price and non-price competitiveness.

## 1 Divergence within the EU: a question of labour costs?

Economic divergence within the EU has inspired a rich literature at the intersection of Comparative Political Economy (CPE) and international macroeconomics. This brief review will focus, in keeping with the focus of the paper, on the trade and competitiveness aspects of the debate.

CPE has largely adopted the growth model approach proposed by Lucio Baccaro and Jonas Pontusson (Baccaro and Pontusson, 2016). The Growth Model Perspective (GMP) aims to bring macroeconomic factors and the role of demand back into CPE. According to the authors, the development of the Varieties of Capitalism theory (Hall and Soskice, 2001) had resulted in a literature too focused on firm-level factors and institutional complementarities. The GMP looks at national accounts aggregates and classifies countries according to what drives their growth: wages, credit, exports, nothing. National economies are thus classified by *growth model*. International trade appears in the GMP first and foremost as a constraint on the regimes available to a given national economy. Domestic demand-led growth is thus only possible if exports have sufficiently high non-price competitiveness (NPC), allowing wages to rise without the trade balance deteriorating. Baccaro and Pontusson measure NPC by estimating the price elasticities of exports – if these are negative and significant, the economy in question must compete on prices and therefore cannot afford wage rises. If, on the contrary, there is no significant price elasticity, the economy benefits from non-price competitiveness, for instance because its goods are of higher quality. Sweden and the United Kingdom are thus found to have good non-price competitiveness, opening the door to a domestic demand-led growth regime, while Germany and Italy must rely on wage moderation (Baccaro and Pontusson, 2016). In the GMP, price competitiveness thus appears, in the spirit of Thirlwall (1979), as a constraint on wage rises and therefore on income growth, while non-price competitiveness allows an economy to escape this constraint.

The respective weights of price and non-price competitiveness are the subject of debate (hereafter the price vs non-price debate) in a literature at the intersection of post-Keynesian macroeconomics and international economics. One side of the debate insists on the role of price competitiveness – and therefore of wage dynamics – in divergence within the EU. This is

the case of Costas Lapavitsas and his co-authors (2012), for whom the euro area crisis is the result of diverging labour costs between countries sharing a common currency, and therefore unable to adjust their external balances through the exchange rate. This diagnosis is shared by European policymakers such as Mario Draghi and Jean-Claude Trichet, successive presidents of the European Central Bank, as well as the then presidents of the European Council and of the European Commission, Donald Tusk and Jean-Claude Juncker (see Storm and Naastepad, 2016). This thesis served to justify the European institutions' exhortations to southern European countries to make their labour markets more flexible in order to bring down wages, an "internal devaluation" destined to make their exports competitive and resolve their external imbalances (Ibid.). Lapavitsas and his co-authors base their argument on comparisons of the evolution of unit labour costs in the countries hit by the crisis (Portugal, Ireland, Italy, Greece, Spain) and in Germany, explaining Germany's superior competitiveness by the stagnation of its ULCs relative to the rising ULCs of its competitors (2012, pp. 22–28).

Proponents of the non-price competitiveness thesis have contested this focus on labour costs in various ways. They argue that since export success has much less to do with relative prices and labour costs than with technological sophistication or quality, diverging economic performance within the EU is fundamentally determined by the different technological capabilities of the economies concerned.

Storm and Naastepad (2015; 2016) thus show that labour costs only represent a minimal part (about 15% in the southern euro area countries) of manufacturing cost structures and should therefore only have a marginal impact on manufacturing prices, and thus on exports. According to these authors, Germany's export success is explained by its higher technological sophistication – it sees stronger productivity growth and gains market share in more technologically advanced products and in services. They also argue that wage cuts would lock the southern countries into this unfavourable specialisation, discouraging the investment that would allow them to raise productivity (Storm and Naastepad, 2015, p. 14).

A similar argument is developed in a series of articles by Jakob Kapeller, Claudius Gräbner-Radkowsch and Philipp Heimberger. They analyse intra-EU divergence as the result of technological polarisation, showing that euro area countries can be grouped into a core and a periphery which diverge in terms of growth, unemployment, trade balances and debt ratios. These groups also differ in their technological capabilities, production in the highest and lowest technology categories being unevenly distributed between them. The core countries' positions in high technology, it is argued, protect them from competition from low-wage emerging economies (Kapeller et al., 2019; Gräbner et al., 2020b; Gräbner-Radkowsch et al., 2022). In an article studying a broader sample of economies, the same authors show that the convergence of poorer economies towards richer ones is conditional on their technological capabilities (Gräbner et al., 2020a). Bottega and Romero (2021) show that exports classified as more technologically advanced have more favourable income elasticities and are less affected by price rises. Finally, Kohler and Stockhammer (2022) question the importance of price competitiveness for growth, remarking that countries whose ULC-based real exchange rates appreciated grew faster than their peers both before and after the great financial crisis. They argue that it is house prices, not either price or non-price competitiveness, that most significantly affected growth during this period.

Proponents of price competitiveness generally provide more sophisticated evidence of a causal link between labour costs and competitiveness: several authors thus show, on the basis of econometric estimations, the impact of labour costs – measured by ULCs or ULC-based REERs – on the export performance of European economies (Boggio and Barbieri, 2017; Paternesi Meloni, 2018; Baccaro and Tober, 2022; Pariboni and Paternesi Meloni, 2022; Keil, 2023). Other analyses find a negative relationship between changes in countries’ market shares in given industries and relative unit labour costs, concluding here too that wages play a role in competitiveness (Ederer and Reschenhofer, 2018; Keil, 2024b). In this paper, these are the papers we will be responding to most directly, investigating the potential biases introduced by some of the methods they employed. Our point of departure will be the systematic review by Sascha Keil (2023) of the various methods used to estimate export equations and especially price elasticities. In his literature review, Keil flags the unreliable results attained in estimating these elasticities and tests, on a sample of the eleven founding countries of the Eurozone over the years 1995-2021, a range of time series methods. In a bid to evaluate the potential difficulties posed by the most widely used price indicators, we will first examine their behaviour in the countries and over the period covered by Keil’s (2023) sample. We will then analyse, using Monte Carlo simulations on fictitious data, the potential biases and limits embedded in several methods used in this literature, both on time series and panel data.

In this section, we have briefly reviewed the literature on economic divergence in the EU, focusing on its treatment of the competitiveness question. In the following sections, we discuss the definitions of price and non-price competitiveness and the conceptual problems associated with price competitiveness indicators.

## 2 Competitiveness and the fundamental problem

The literature just surveyed rarely precisely defines competitiveness. The articles by Walter Paternesi Meloni (2021) and Sascha Keil (2024a) which squarely focus on the price vs non-price competitiveness question only implicitly define competitiveness as that which determines export volumes (respectively pages 6 and 1). The two authors define price competitiveness as what is “habitually captured” or “expressed” by REERs or ULCs (Ibid.). Non-price competitiveness’ definition is a bit more fleshed out: it is what “captures the quality and sophistication of exported goods” (Paternesi Meloni, 2021, p. 7) or, in more negative terms, “the sum of all factors other than prices and cost that impact on trade performance” (Keil, 2024a, p. 1). These factors are conceptualised as independent, especially in the latter, negative definition.

In this paper, we will largely follow these definitions and define competitiveness as that which allows a producer to maintain or increase their market share *against or at the expense of another*. The distinction between price and non-price is also rather intuitive – goods are distinguished from each other not just by their prices but also by their characteristics such as ergonomics, functions, tastes, aesthetics. The choice between similar goods or services – two laptops, restaurants, dishes in a restaurant – is determined not just by prices but also characteristics *which can justify a higher price*. Quality is therefore to some extent a substitute for price. Because it is by definition not quantifiable, it complicates the measurement not just of non-price but also of price competitiveness, given that the effects of prices on competitiveness can be attenuated or indeed

more than compensated by quality. Price is only relevant for comparable qualities, and only entirely determinant for identical ones. This will become particularly relevant in the case of a central price, that of labour.

This varying relevance of price poses what we term the fundamental problem in the measurement of competitiveness – the impossibility of measuring the relevant relative prices. Because economies produce a variety of different products, a product-by-product comparison of prices of identical or substitutable products is impossible. Even such a comparison would not necessarily be enlightening, as trade leads to the specialisation of economies whose insertion into global value chains leads them to produce and export different goods. Non-price competitiveness is partly meant to address exactly these differences: it is often associated with technical sophistication or complexity, themselves associated with higher per-worker value added.

It is therefore impossible to directly compare levels of relative prices. How, then, can one go about measuring price competitiveness? Though this framing is, to our knowledge, absent from the literature, we argue that the price measures employed represent two implicit responses to this problem: comparing price changes and comparing a single key price, that of labour.

### 3 Price indexes: measuring price changes

The first solution is adopted by authors who use REERs deflated by price indexes such as consumer prices, producer prices or Gross Domestic Product (GDP) deflators. According to this approach, the prices of exports are affected by nominal exchange rates and the inflation rates affecting prices on either side of the exchange rate. A generalised rise in prices in a given country should thus impact, through the rise in wages and input prices, the relevant prices.

As summarised by Giordano and Zollino (2016, p. 5), the different indexes each have their limits as they either exclude prices that should be included (exports, intermediate goods or services that enter into exporting firms’ costs) or include ones that should not (non-tradables, domestically consumed imports). As such, this first solution falls prey to the fundamental problem: the volume and variety of products and prices which may or may not matter for export performance.

Export price indexes, potentially the closest thing to an index of the relevant prices, pose a similar set of issues for our purposes of measuring domestic cost factors: they are biased by pricing to market (which reflects demand/market conditions rather than costs) and are affected by imported inputs which matter for price competitiveness but are independent from the domestic component we are interested in (Amiti et al., 2013).

The impossibility of isolating within a single index the relevant prices would not be a problem if one could suppose that they moved in concert with the general price level measured by these various price indexes. Unfortunately, one can’t: given the Baumol effect (Baumol and Bowen, 1965), prices follow different paths depending on the underlying productivity dynamics, with the sectors experiencing faster productivity growth exhibiting falling relative prices. Given faster productivity gains in manufacturing, which makes up most of international trade, one would expect the movements of export prices to be largely decorrelated from general inflation. Indeed, this effect has its international counterpart in the Balassa-Samuelson effect (Balassa,

1964; Samuelson, 1964) resulting in higher overall price levels in economies with more productive exporting sectors.

These mechanisms constitute reasons to doubt that the domestic cost components of exported goods follow the general price level. They do not imply, however, that there are *no* correlations between the relevant prices and prices in general – exporting and non-exporting activities remain linked by common inputs, notably labour, which can induce movements in the same direction. Covariation across price indices and general tendencies in price levels across nine EU countries in the period 1995-2021 are summarised in Table 1. The table clearly shows the Baumol effect, with manufacturing deflators rising about three times less than their GDP equivalents and the cumulative rises being correlated with an  $R^2$  of only 0.5 across countries. The year-on-year changes are even more weakly correlated, with an  $R^2$  of only 0.18 – general price movements are a weak predictor of manufacturing price movement.

Goods export prices rise slightly faster on average than manufacturing prices, which may have to do with the price dynamics of imported inputs or, more likely, the non-manufacturing inputs entering exported goods. Cumulative export price rises are more correlated with rises in the manufacturing value added deflator than with GDP deflator movement. On a year-on-year basis, export price movements are almost completely decorrelated from any of the other price indices.

Table 1: Price index rises and correlations, Eurozone founding members minus Ireland, 1995–2021

| Index                      | Mean cumulative<br>rise, 1995–2021 (%) | Cumulative rises vs |                  | Yearly changes vs |                  |
|----------------------------|----------------------------------------|---------------------|------------------|-------------------|------------------|
|                            |                                        | GDP deflator        | Mfg. VA deflator | GDP deflator      | Mfg. VA deflator |
| GDP deflator               | 61.8                                   | –                   |                  | –                 |                  |
| Manufacturing GVA deflator | 20.5                                   | 0.49 (0.33)         | –                | 0.18 (0.21)       | –                |
| Export deflator, goods     | 27.0                                   | 0.27 (0.36)         | 0.68 (1.20)      | 0.03 (0.05)       | 0.01 (0.07)      |

*Notes:* Mean cumulative rises in %, averaged across countries. The  $R^2$  and slopes are computed on log changes. Cells report the  $R^2$ , with the slope in brackets; the slopes are simple OLS coefficients of the column index regressed on the row index. The correlations of cumulative rises are across countries, the correlations of year-on-year rises are pooled over countries and years. The means exclude Ireland, the correlations exclude Ireland and Luxembourg, respectively due to the bias introduced by multinationals' profit-shifting and the weight of financial services in exports breaking the link between export and manufacturing deflators. Data sources: AMECO (GDP and manufacturing value-added deflators); Eurostat, national accounts (implicit deflator of exports of goods).

This analysis of the movements of price indexes illustrates the fundamental problem: the diversity of productive activities, underlying cost structures, productivity and market dynamics means these indices are at best quite rough measures of the costs relevant to competitiveness. Prices in general diverge from manufacturing prices due to the Baumol effect and export prices rise cumulatively less than overall prices but more than the manufacturing deflators, with their own index capturing elements outside of domestic cost pressures. The conducted regressions suggest that price movements are especially orthogonal over short periods, with year-on-year changes barely correlated. The price movements relevant to the cost structures of exporting firms are therefore all the less likely to be captured in estimations of short-term relationships. Finally, the regression coefficients of one index on another are all quite far from unity, suggesting estimated elasticities should vary widely across indices.

Given these a priori reasons to distrust price indexes as a measure of the costs relevant to the

price competitiveness of exporting firms, the not infrequent negative correlation of price index-deflated REERs with exports is somewhat puzzling. Do these results not indicate that price indices do after all track what they are supposed to, i.e. exporters' costs? On the face of the evidence presented, this is plausible for longer-term relationships: the cumulative correlations between the various indices are far from negligible, suggesting enough covariations for indices to somewhat track exporters' costs. The significant long-run and especially short-run relationships detected could also, however, be explained by the exchange rate component of REERs rather than the deflators.

Whatever the explanation, the evidence and arguments presented should inspire caution when choosing the deflators to use in export equations and interpreting the results, especially in estimations of short-run relationships.

## 4 Unit labour costs: measuring a key price?

The second solution to the fundamental problem is to use a key price, that of labour. This solution is implicitly adopted when using ULC-deflated REERs. It is on its face quite seductive: the classical breakdown of relative prices originating in Adam Smith's *Wealth of Nations* and continued in classical economists from David Ricardo through Karl Marx to Anwar Shaikh (Shaikh, 2016, pp. 380–388) demonstrates that all relative prices can be decomposed into a product of vertically integrated unit labour costs and profit-wage ratios. This solution thus promises to bypass the fundamental problem by comparing the price that determines (along with productivity) all other prices. Because it focuses on a single commodity, it in principle allows for the comparison of price levels, not just price changes as when using price indices. These advantages are however in practice questionable.

Firstly, levels are seldom compared: most ULC-based comparisons use indices normalised to a certain year, eliminating this potential advantage. This practice has its theoretical justification – ULCs represent different fractions of the cost structure of different producers, meaning levels comparisons would not necessarily be relevant.

Secondly, labour is arguably not a homogeneous commodity: wage levels differ wildly across countries, and while imperfect labour mobility certainly plays a role, these disparities also certainly reflect differences in skills and productivity. The use of ULCs rather than wage levels addresses this difficulty by dividing wages by productivity, controlling for these differences in skills and yielding a wage cost per unit of value added.

The introduction of productivity unfortunately means ULCs are not, as their name suggests, a straightforward measure of cost – their evolution depends not just on wage movements, but also on changes in productivity. Their rise can therefore result not only from a rise in wages – as implied when taking them to be a measure of wage costs – but also from a fall in productivity. Reciprocally, a fall in a country's relative ULC can indicate successful wage restraint but also favourable productivity dynamics. An additional complexity is introduced by the fact that *nominal* ULC, those most frequently used, are built by dividing nominal wages by real productivity, meaning a deflator is introduced into the mix. As pointed out by Felipe and Kumar (2014), this means they are effectively a wage share multiplied by a price index, either the GDP deflator or the manufacturing value-added deflator. Before discussing the difficulties introduced by the

price index component, we will examine those relative to the wage share component.

Productivity being value added divided by employment, it is potentially procyclical, tending to fall during crises, not because workers become less productive in any real sense, but because falling sales tend to only be followed by reductions in employment with a certain lag (Biddle, 2014; Dossche et al., 2023). Any lag in the adjustment of the wage bill to sales – because of either labour hoarding or wages lagging productivity – will in fact mechanically reduce ULCs in prosperous sectors or countries and increase them in struggling ones. For this effect not to be observed, firms would have to hire or increase wages as fast or faster than their sales increase, or fire as fast or faster than their sales decline. Rising ULCs can thus be an expression rather than a cause of a sector or an economy’s difficulties. Estimating regressions of economic success – export or value-added growth – on ULCs or ULC-deflated REERs therefore suffers from potential endogeneity. As shown by Figure 1, the short-run covariation between exports and value added – ULCs’ denominator – is very strong, with an  $R^2$  of 0.66 for our sample and period.

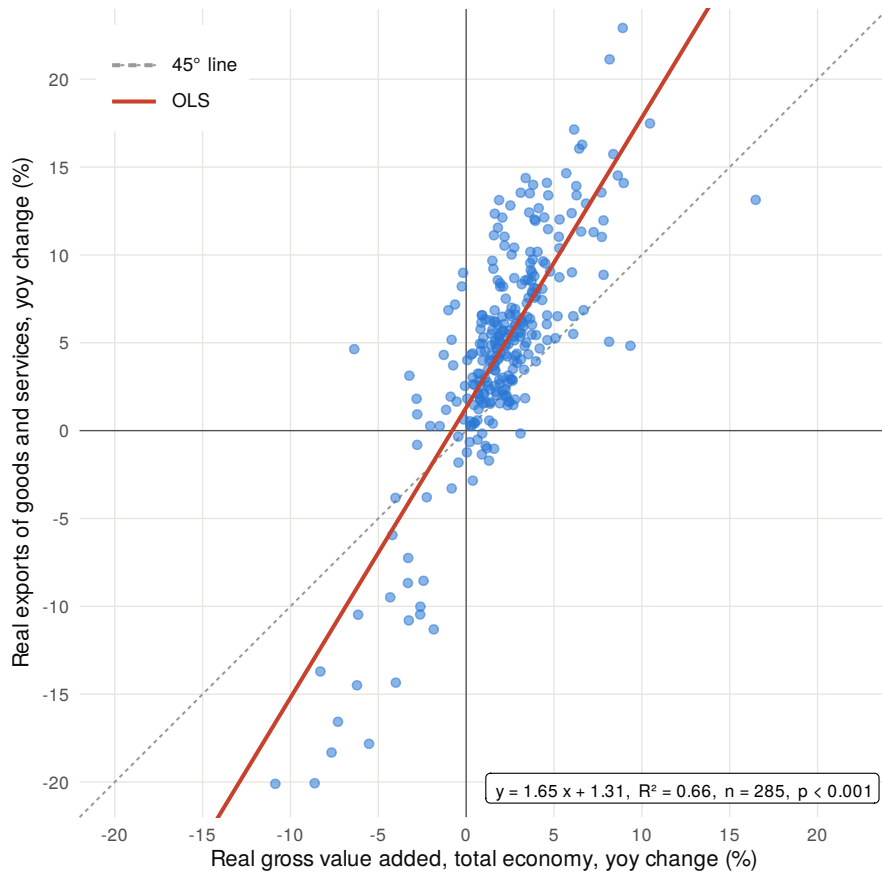


Figure 1: Real value added and real exports, yearly changes, 1996–2021

*Notes:* Year-on-year % changes, eleven euro area founding members, one point per country-year (285 observations; Ireland 2015 excluded because of the relocation of intangible assets). Same scale on both axes. Source: AMECO, gross value added at 2020 prices (total of branches), exports of goods and services at constant prices.

An additional reason makes ULCs a poor measure of firms’ costs: heterogeneity in cost structures. As pointed out by Storm and Naastepad (2016), labour costs only represent a small fraction of manufacturing firms’ prices, about 15% in southern Europe and 24% in Germany. To the extent that firms’ prices on average reflect total costs plus a given margin, the wage

share should be inversely proportional to their capital intensity (organic composition of capital, if you will). Higher ULCs would in that case only indicate lower competitiveness if the latter were a function of lower proportional input costs or fixed costs. As explained by Storm and Naastepad (Ibid.), this should result in fairly low wage elasticities of exports, as the wage elasticities of prices are diluted by wages' low participation in total costs. This feature of ULCs also means that changes in ULCs can reflect outsourcing (either outside of the manufacturing sector, if using manufacturing ULCs, or outside of the country) or changes in production technique, which certainly can contribute to competitiveness, but not through the conventional "price competitiveness" channel of wages' pressure on costs.

These features of the wage share component of ULCs constitute good reasons to doubt that they – and by extension ULC-deflated REERs – reliably reflect the wage component of exporting firms' costs: they are tainted by the procyclicality and endogeneity of productivity, present an incomplete picture of exporters' cost structures and can be affected by outsourcing and internationalisation.

When nominal ULCs are used, additional difficulties are introduced by their price index component. Given the stability of the wage share relative to price indices, ULCs are in fact mainly driven by the latter component, as shown by Felipe and Kumar (2014) for twelve European economies for 1980-2007 and illustrated for our period by Figure 2. In the long term, ULC is therefore practically a price index. In the short term, however, the wage share accounts for most of its movement, the wage share's volatility in turn being very slightly dominated by the value-added component.

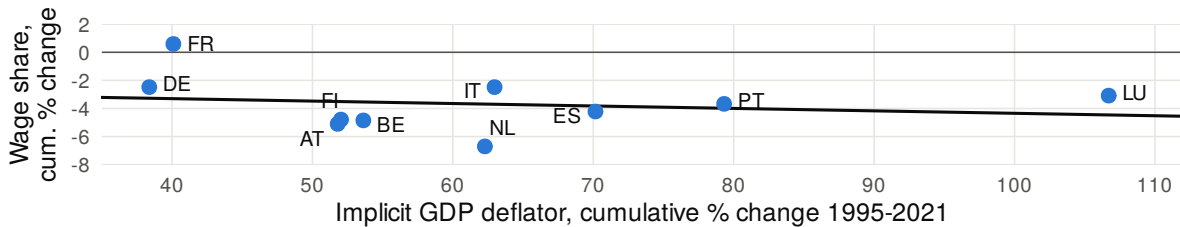


Figure 2: Wage share and GDP deflator, cumulative % change 1995–2021

*Notes:* Decomposition of nominal ULCs into the wage share (real ULC) and the implicit GDP deflator (Felipe and Kumar, 2014), one point per country, total economy. Same scale on both axes; solid line: OLS fit (slope  $-0.02$ , s.e.  $0.03$ ). Ireland not shown (wage share  $-45\%$ , deflator  $+79\%$ ). Source: AMECO, nominal and real unit labour costs, total economy.

This carries over to the decomposition of the changes in ULC-deflated REERs: as shown in Figure 3, the cumulative change is driven mainly by the relative GDP deflator, with the nominal exchange rate showing the second largest contribution, except in the cases of Italy and Spain where nominal appreciation dominates. The year-on-year variance is driven mainly by the nominal exchange rate, which accounts for over 50% of the variance in 8 of the 11 countries.

Three conclusions can be drawn from this decomposition of ULCs' and ULC-deflated REERs'

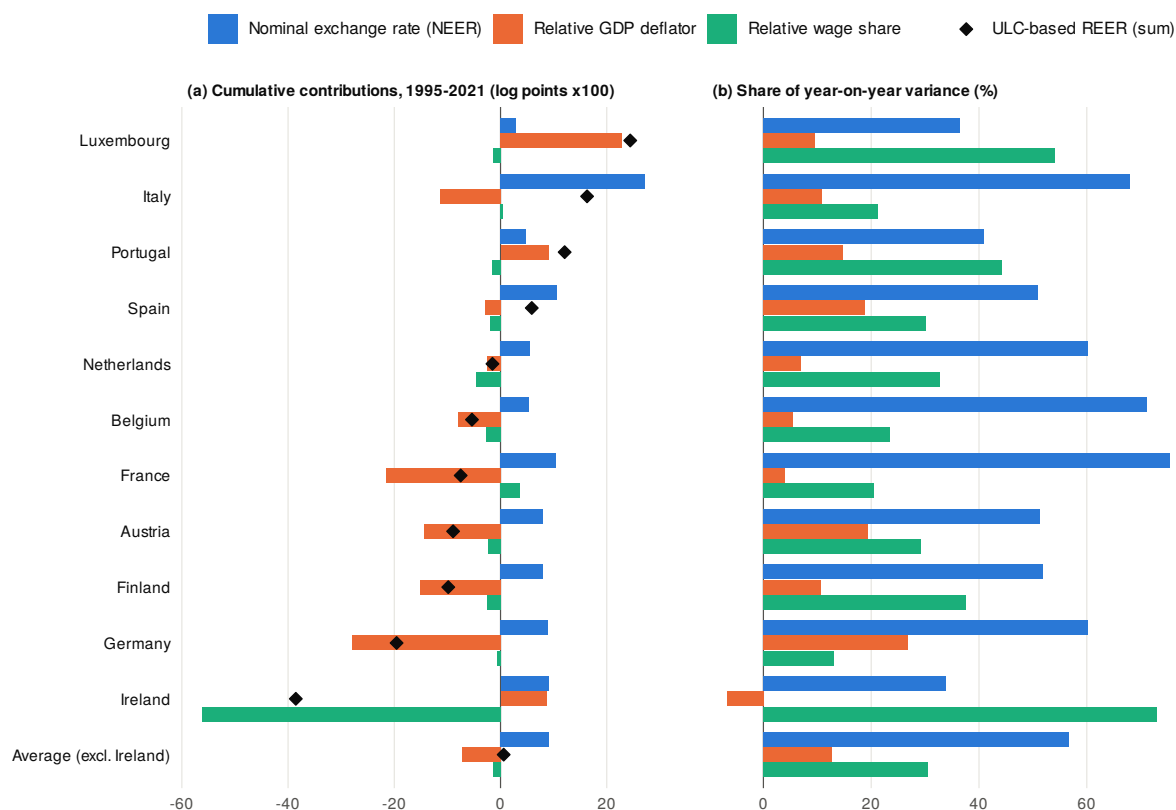


Figure 3: Decomposition of the ULC-based REER, 1995–2021

*Notes:* The log of the ULC-based REER is split into the nominal effective exchange rate (NEER), the relative GDP deflator (GDP-deflator REER / NEER) and the relative wage share (ULC-REER / GDP-deflator REER). (a) Contributions to the cumulative change, in log points ( $\times 100$ ; positive = real appreciation); the components add up to the total (diamond). Countries sorted by the total change. (b) Contribution of each component to the variance of the year-on-year change in the ULC-REER,  $\text{cov}(\Delta \text{component}, \Delta \text{REER}) / \text{var}(\Delta \text{REER})$ ; shares add up to 100%. The average excludes Ireland. Source: European Commission, price and cost competitiveness database, 37 trading partners.

variations. First, in the long term, ULCs are practically a price index. Second, in the short term, they are mainly driven by their wage share component, which is countercyclical and thus potentially endogenous to economic performance. Third, ULC-REERs are likewise over the longer term driven by the price and exchange rates components, while the wage share can play a role in the short run. In the context of standard export equations, these facts should indicate that ULCs and ULC-REERs can allow for the identification of the effects of wages on competitiveness but also produce spurious results due to endogeneity in the short run, while they are less likely to exhibit either this advantage or this problem when analysing long-run relationships.

## 5 Humbug labour cost competitiveness?

In the previous section, we discussed the composition of ULCs, ULC-REERs and their movement, laying out some hypotheses regarding what the variables' behaviour means for the identification of relationships between these variables and exports. In this section, we reproduce some of the econometric methods used in the literature on fictitious data to ascertain the extent of the endogeneity problem and what underlying patterns these methods can or cannot detect.

The methods tested are (i) OLS on differences, (ii) ARDL estimation with bounds tests (Pesaran et al., 2001), (iii) panel fixed effects and (iv) mean group panel ARDL. OLS on differences and ARDL with bounds tests are chosen following Sascha Keil's (2023) systematic review of various time series methods for the estimation of trade elasticities. Keil selects OLS on differences with the Prais-Winsten AR(1) estimation method as the problematic default baseline, and finds ARDL models to be preferable, hence our retaining them to test where the problems posed by ULCs and ULC-REERs can or cannot create bias or spurious results. Panel fixed effects, as used by Ederer and Reschenhofer (2018), and panel mean group ARDL as used by Keil in a subsequent paper (2024b), are chosen as representative of the subset of the literature that estimates sector-level ULC effects.

OLS on differences estimates the following equation:

$$\Delta x_t = c + \beta \Delta f d_t + b \Delta reer_t + u_t, \quad u_t = \rho u_{t-1} + \varepsilon_t \quad (1)$$

where  $x_t$ ,  $f d_t$  and  $reer_t$  are the logs of real exports, foreign demand and the ULC-REER, and  $b$  is the (short-run) price elasticity. The residuals are allowed to follow an AR(1) process: the equation is estimated by feasible generalised least squares with the Prais-Winsten transformation, iterating on  $\rho$  until convergence, which is the baseline specification in Keil (2023).

The ARDL bounds-test estimates the following error correction model:

$$\Delta x_t = c + a x_{t-1} + b_1 f d_{t-1} + b_2 reer_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-i} + \sum_{j=0}^{q-1} (\delta_{1j} \Delta f d_{t-j} + \delta_{2j} \Delta reer_{t-j}) + e_t \quad (2)$$

The coefficient  $a$  is the error correction (adjustment) term and the long-run elasticities are  $\theta_{fd} = -b_1/a$  and  $\theta = -b_2/a$ , with standard errors obtained by the delta method. In the ARDL( $p, q_1, q_2$ ) notation used below,  $p$  is the number of lags of the dependent variable and  $q_1, q_2$  those of the regressors in the levels representation of the model, which translates into the

differences  $\Delta_t$  to  $\Delta_{t-q+1}$  in equation (2): an ARDL(1,1,1) thus includes the contemporaneous differences of the regressors but no lagged differences. The lag orders are chosen between 1 and 4 by the Akaike information criterion, on a common sample, with the same lag order  $q$  for both regressors. The bounds test (Pesaran et al., 2001) tests the null of no levels relationship ( $a = b_1 = b_2 = 0$ ) with an  $F$ -test, completed by a  $t$ -test on  $a$ . We use the asymptotic 5% critical values for case III (unrestricted intercept, no trend) with  $k = 2$  regressors and consider a levels relationship detected when both statistics lie beyond the upper, I(1), bounds ( $F > 4.85$  and  $t < -3.53$ ).

The panel fixed effects estimation uses the following equation:

$$\Delta vas_{it} = \alpha_i + \lambda_t + \beta \Delta rulc_{it} + \varepsilon_{it} \quad (3)$$

Finally, the panel mean group ARDL estimates the following equation:

$$\Delta vas_{it} = c_i + \alpha_i vas_{i,t-1} + \beta_i rgdp_{i,t-1} + h_i rulc_{i,t-1} + \Gamma'_i \Delta Z_{it} + e_{it} \quad (4)$$

where  $rgdp_{it}$  is the log of the country's share of total domestic demand (relative GDP) and  $\Delta Z_{it}$  the short-run terms. We consider two specifications: an ARDL(1,1,1), with  $\Delta Z_{it} = (\Delta rgdp_{it}, \Delta rulc_{it})$ , the specification most commonly estimated, and the specification of Keil (2024b, eq. 2),  $\Delta Z_{it} = (\Delta vas_{i,t-1}, \Delta rgdp_{i,t-1}, \Delta rulc_{i,t-1})$ , which includes lagged differences only and corresponds to an ARDL(2,2,2) without contemporaneous regressors.

The mean group estimator (Pesaran and Smith, 1995) estimates equation (4) separately for each country-sector and averages the unit-level long-run coefficients:  $\theta_i = -h_i/\alpha_i$ ,  $\hat{\theta}_{MG} = N^{-1} \sum_i \hat{\theta}_i$ , with standard error  $s_\theta/\sqrt{N}$ , where  $s_\theta$  is the cross-unit standard deviation of the  $\hat{\theta}_i$ . Because  $\theta_i$  is a ratio, units with an adjustment coefficient  $\alpha_i$  close to zero produce extreme estimates of either sign; we therefore also report a variant with the mean group trimmed at 5% in each tail, its standard error computed from the winsorised standard deviation. Significance of the mean group adjustment coefficient  $\hat{\alpha}_{MG}$  plays the role of the bounds test as evidence of a levels relationship.

The first data generating process, DGP1, is constructed specifically to ascertain the extent of the endogeneity problem. It is given by equations (5)–(16) and its calibration is detailed in Annex A1: exports are a function of foreign demand and persistent exogenous shocks, with no role for the REER, foreign demand is a random walk with a drift, as is domestic demand. Value added is a weighted sum of exports and domestic demand, target employment is a function of value added and trend productivity, to which employment adjusts with a lag (the labour hoarding mechanism). Trend productivity is exogenous, while measured productivity is given by value added divided by employment, which the labour-hoarding mechanism makes procyclical. Wages are entirely exogenous, with a trend equal to that of productivity and random persistent shocks making the wage share I(1), while the REER is a product of the wage share and an exogenous random walk standing in for partners' ULCs and the nominal effective exchange rate. The simulation runs over 107 periods to match the 26-year period covered by Keil's (2023) sample. DGP1 abstracts from technology shocks and from feedback from exports to domestic demand. Adding stochastic trend productivity leaves the results essentially unchanged, while a

multiplier would reinforce the endogeneity, so its omission is conservative.

$$fd_t = fd_{t-1} + \mu_{fd} + \varepsilon_t^{fd} \quad (5)$$

$$v_t = \rho_v v_{t-1} + \varepsilon_t^v \quad (6)$$

$$x_t = \beta_{fd} fd_t + v_t \quad (7)$$

$$dd_t = dd_{t-1} + \mu_{dd} + \varepsilon_t^{dd} \quad (8)$$

$$y_t = s_x x_t + (1 - s_x) dd_t \quad (9)$$

$$a_t = a_0 + g_a t \quad (10)$$

$$l_t = l_{t-1} + \varphi (y_t - a_t - l_{t-1}) + \varepsilon_t^l \quad (11)$$

$$q_t = y_t - l_t \quad (12)$$

$$w_t = w_{t-1} + g_a + \varepsilon_t^w \quad (13)$$

$$ws_t = w_t - q_t \quad (14)$$

$$z_t = z_{t-1} + \varepsilon_t^z \quad (15)$$

$$reer_t = ws_t + z_t \quad (16)$$

All variables are in logs and all shocks are independent and normally distributed. Exports  $x_t$  depend on foreign demand  $fd_t$ , with an income elasticity  $\beta_{fd}$ , and on an export-specific shock  $v_t$  – non-price competitiveness, if you will – which is persistent but stationary for  $\rho_v < 1$ . Value added  $y_t$  is a weighted sum of exports and domestic demand  $dd_t$ . Employment  $l_t$  adjusts towards its target  $y_t - a_t$ , given trend productivity  $a_t$ , at speed  $\varphi$ : with  $\varphi = 1$  employment adjusts within the period and measured productivity  $q_t$  is trend productivity up to the employment noise, while  $\varphi < 1$  is the labour-hoarding case. Wages  $w_t$  follow a random walk with a drift equal to trend productivity growth, so that  $ws_t$ , the wage share (the ULC of the DGP, prices being normalised), is I(1) and, when  $\varphi = 1$ , independent of exports. The REER adds to it  $z_t$ , an exogenous random walk standing in for the relative price deflator and the nominal effective exchange rate; with  $\sigma_z = 0$  the REER is the wage share alone.

We also constructed a panel version of the DGP1 to test the methods used in papers testing sectoral ULCs' effects on countries' shares of a given sector's value added (Ederer and Reschenhofer, 2018; Keil, 2024b). DGP1p is functionally similar to the non-panel version, with shock sizes and adjustment parameters calibrated to realistic magnitudes for annual variations. Its details are in Annex A2.

$$Y_{cs,t} = Y_{cs,0} \left[ (1 - x_{cs}) \frac{d_{c,t}}{d_{c,0}} + x_{cs} \frac{D_{s,t}}{D_{s,0}} \right] \exp(\theta rulc_{cs,t} + v_{cs,t}) \quad (17)$$

$$v_{cs,t} = \rho_v v_{cs,t-1} + \varepsilon_{cs,t}^v \quad (18)$$

$$vas_{cs,t} = \ln Y_{cs,t} - \ln \sum_{c'} Y_{c's,t}, \quad rulc_{cs,t} = ulc_{cs,t} - \frac{1}{N_c} \sum_{c'} ulc_{c's,t} \quad (19)$$

Real value added of sector  $s$  in country  $c$  starts from the country's GDP times the sector's share in European GDP,  $Y_{cs,0}$ , and then follows a weighted average of country domestic demand

$d_{c,t}$  and European demand for the sector  $D_{s,t}$  (both random walks with drift), the weight  $x_{cs}$  being the export share of the country-sector’s demand. The export shares differ across sectors, following the export dependence of manufacturing sectors reported by Keil (2024b, Table 1), and across countries, through a log-normal openness factor, so that a common European sector shock moves countries’ value-added shares. Employment, productivity and wages follow equations analogous to those of DGP1 at the unit level, with wage shocks made of a country and a country-sector component. The relative GDP regressor of the panel ARDL,  $rgdp_{c,t}$ , is generated from the same domestic demand series, as the log of the country’s share of total domestic demand. With  $\theta \neq 0$ , value added and measured productivity are solved jointly within the year. The calibration is detailed in Annex A2.

First, DGP1 was run using a Monte Carlo simulation for different specifications, to test our first two econometric methods, OLS on differences and ARDL with bounds test. Table 2 reports the percentage of estimations that return a false positive as a function of the strength of the labour hoarding mechanism. For OLS on differences, what counts as a false positive is a statistically significant (at 5%) coefficient for the REER, which is by construction independent of exports. Similarly, for the ARDL, a false positive means a statistically significant long-run REER coefficient and bounds test scores past the upper bound critical values. Table 2 also reports the false positive proportions for a simulation over twice as many quarters ( $T=214$ ) and differentiates between cases when the ULC-REER is just the wage share and when it contains the exogenous “price” (equivalent to estimating on the wage share vs ULC-REER).

Table 2: Results of methods (i) and (ii) estimated on DGP1 runs, 1’000 runs per cell

| $\sigma_z$ | $T$ | (i) OLS on differences |               | (ii) ARDL with bounds test |                  |             |                   |
|------------|-----|------------------------|---------------|----------------------------|------------------|-------------|-------------------|
|            |     | mean $b$               | % sig. neg.   | median $\theta$            | % bounds rejects | % sig. neg. | % false positives |
| 0          | 107 | 0.00 / -1.23 / -1.38   | 2 / 100 / 100 | -0.01 / -0.92 / -1.37      | 38 / 18 / 7      | 8 / 20 / 29 | 6 / 8 / 5         |
| 0          | 214 | 0.00 / -1.23 / -1.40   | 3 / 100 / 100 | 0.00 / -0.51 / -1.03       | 89 / 46 / 15     | 5 / 17 / 29 | 5 / 10 / 8        |
| 0.015      | 107 | 0.00 / -0.27 / -0.40   | 3 / 63 / 95   | 0.00 / -0.12 / -0.24       | 39 / 33 / 27     | 9 / 11 / 18 | 6 / 7 / 10        |
| 0.015      | 214 | 0.00 / -0.28 / -0.40   | 2 / 92 / 100  | 0.00 / -0.08 / -0.12       | 88 / 83 / 70     | 4 / 13 / 16 | 4 / 12 / 14       |

*Notes:* Each cell shows three values, one per employment adjustment speed  $\varphi$ , in the order 1 / 0.5 / 0.2 ( $\varphi = 1$ : no labour hoarding). True REER elasticity of exports: 0.  $\sigma_z$ : standard deviation of the shocks to the exogenous component of the REER ( $\sigma_z = 0$ : the REER is the wage share alone).  $T$ : number of quarters. % sig. neg.: share of runs with a REER coefficient negative and significant at 5%. % bounds rejects: share of runs in which the bounds test detects a levels relationship. % false positives: share of runs with both a detected levels relationship and a significant negative long-run REER coefficient.

Over the shorter sample, method (i) is very vulnerable to the endogeneity of ULCs to exports, reporting false positives in all runs with any labour hoarding when estimating on ULCs and a majority of runs for a  $\varphi$  of 0.5, and nearly all runs for a  $\varphi$  of 0.2 when using ULC-REERs. Method (ii) is much more robust to short-run endogeneity, regardless of the independent variable used, though it still produces false positives at more than the nominal rate, especially in the longer sample.

OLS on differences is thus particularly vulnerable to the short-run endogeneity of the wage share to value added, into which exports enter, identifying a significant “price” coefficient even where no price effects exist. The level of false positives is slightly lower when using ULC-REERs as the noise from the exogenous price term somewhat weakens the link through the ULC denominator. The ARDL, though also biased, is much more robust in terms of false positives

because it is designed to detect long-run co-movement and the endogeneity is purely a short-run phenomenon, with export deviations from trend and ULC deviations mirroring each other. With the export-driven gap between target and actual employment – which drives the wage share’s endogeneity to exports – reverting to 0, there is no long-run relationship to detect.

Table 3: Results of methods (iii) and (iv) run on DGP1p runs, 500 runs per cell

| Wage noise | $T$ | ARDL(2,2,2)* |        | ARDL(1,1,1)  |        | FE growth    |         |
|------------|-----|--------------|--------|--------------|--------|--------------|---------|
|            |     | coef.        | % sig. | coef.        | % sig. | coef.        | % sig.  |
| ×1         | 24  | -0.01 / 0.04 | 2 / 1  | 0.00 / -1.24 | 1 / 42 | 0.00 / -0.98 | 3 / 100 |
| ×1         | 15  | 0.01 / -0.06 | 1 / 3  | 0.00 / -1.16 | 2 / 44 | 0.00 / -0.98 | 3 / 100 |
| ×3         | 24  | 0.00 / 0.02  | 3 / 0  | 0.00 / -0.26 | 3 / 85 | 0.00 / -0.44 | 2 / 100 |
| ×3         | 15  | 0.00 / 0.00  | 2 / 1  | 0.00 / -0.32 | 1 / 62 | 0.00 / -0.45 | 3 / 100 |

*Notes:* Each cell shows no endogeneity ( $\varphi = 1$ ) / labour hoarding ( $\varphi = 0.3$ ). True relative ULC elasticity: 0. Coefficients are medians across runs; % sig.: share of runs with a coefficient negative and significant at 5%. Wage noise: standard deviation of the wage shocks relative to the baseline calibration (Annex A2). The ARDLs are run on 28 countries times 18 sectors, for the number of years indicated in the second column, with the (untrimmed) mean group estimator. The fixed effects estimation is conducted on a panel of 132 randomly drawn country-sectors, for  $T = 24$  in the first and third rows and  $T = 13$  in the second and fourth. \*ARDL(2,2,2) is Keil’s (2024b) equation 2, with no contemporaneous regressors.

Table 3 reports the results of methods (iii) and (iv) run on DGP1p simulations. The spurious results of the OLS on differences also affect the panel regression of growth of value-added shares on relative ULC change. In the panel version of our DGP with no true cost effect, the fixed effects estimator yields a RULC coefficient, significant in every run, of between -0.4 and -1 depending on the share of wage shocks in relative ULC variation (Ederer and Reschenhofer’s (2018) estimates vary from -0.22 to -0.38)<sup>1</sup>.

While the ARDL method seemed quite robust to the endogeneity problem for single-country export equations, this is not the case for panel ARDL with the mean group estimator, depending on the lag specification. With endogeneity, an ARDL(1,1,1) yields significant spurious results for between 42 and 85% of the runs, depending on the run’s length and the amount of wage noise in the REER. With the mean group trimmed at 5% in each tail, the test has correct size without endogeneity, and the spurious hoarding elasticity is significant in all replications. The specification with no contemporaneous regressors Keil (2024b, p. 202) uses, on the other hand, produces no spurious results<sup>2</sup>. The mean group panel ARDL’s vulnerability to the endogeneity bias is thus real but depends crucially on the specifications.

For the endogeneity of ULCs to value-added – and thereby exports – to reliably produce spurious results in single-country ARDL estimations, exogenous export performance would have to *durably* move the wage share, e.g. with a permanent change to productivity from a Verdoorn effect without proportional adjustment in wages or a very strong labour hoarding effect. What this also implies, however, is that the wage share component of ULCs would have to exhibit persistent autonomous variation for an ARDL estimation to detect its independent effects on exports.

<sup>1</sup>We did not check whether their system GMM estimation, with lagged independent variables, fixes the endogeneity problem as they suggest it should. Given persistent export shocks, as in our DGP, it might not.

<sup>2</sup>Though it does attenuate real same year effects: in runs with a true price elasticity, a true elasticity of -0.5 came out at -0.2 to -0.35.

To test these twin possibilities, we modify the data generating process to yield DGP2. It is identical to DGP1 except for the following two points. First, productivity is not just a trend but also has a Verdoorn effect: it is durably affected by output growth, with above trend output growth raising productivity and vice versa. Second, there is a “price” component to the export equation, with exports a function of foreign demand, exogenous shocks and ULC-REER (the wage share with or without an added random walk to stand in for the other components).

$$a_t = a_{t-1} + g_a + \beta_V (\Delta y_t - \bar{g}_y) \quad (20)$$

$$x_t = \beta_{fd} f d_t + \theta reer_t + v_t \quad (21)$$

where  $\bar{g}_y$  is the expected growth rate of value added and  $\beta_V$  the Verdoorn coefficient. Because exports depend on the contemporaneous REER, which depends on measured productivity and hence on value added, exports, value added and productivity are solved jointly within each period.

Table 4: Results of methods (i) and (ii) estimated on DGP2 runs, 1’000 runs per cell

| $\sigma_z$ | True $\theta$ | (i) OLS on differences |                 | (ii) ARDL with bounds test |                  |              |              |
|------------|---------------|------------------------|-----------------|----------------------------|------------------|--------------|--------------|
|            |               | median $b$             | % sig. neg.     | median $\theta$            | % bounds rejects | % sig. neg.  | % both       |
| 0          | 0             | -1.40 / -1.40 / -1.41  | 100 / 100 / 100 | -1.48 / -1.21 / -0.99      | 9 / 13 / 15      | 34 / 41 / 42 | 7 / 10 / 10  |
| 0          | -0.3          | -1.55 / -1.58 / -1.58  | 100 / 100 / 100 | -1.70 / -1.46 / -1.27      | 7 / 11 / 14      | 44 / 52 / 61 | 6 / 9 / 13   |
| 0          | -1            | -1.96 / -1.96 / -1.96  | 100 / 100 / 100 | -2.49 / -2.17 / -1.85      | 9 / 12 / 13      | 73 / 83 / 89 | 9 / 12 / 13  |
| 0.015      | 0             | -0.40 / -0.42 / -0.43  | 96 / 97 / 98    | -0.19 / -0.23 / -0.19      | 27 / 29 / 31     | 16 / 22 / 20 | 8 / 12 / 11  |
| 0.015      | -0.3          | -0.67 / -0.68 / -0.69  | 100 / 100 / 100 | -0.53 / -0.52 / -0.50      | 25 / 26 / 31     | 51 / 55 / 56 | 20 / 19 / 24 |
| 0.015      | -1            | -1.28 / -1.29 / -1.29  | 100 / 100 / 100 | -1.22 / -1.19 / -1.19      | 26 / 27 / 31     | 95 / 98 / 97 | 26 / 27 / 31 |

*Notes:* Each cell shows results for a Verdoorn coefficient  $\beta_V$  of 0 / 0.3 / 0.6. Labour hoarding with  $\varphi = 0.2$ , persistence of the export shocks  $\rho_v = 0.85$ ,  $T = 107$  quarters. % both: share of runs with both a detected levels relationship and a significant negative long-run REER coefficient. Other definitions as in Table 2.

DGP2 is run with 1’000 iterations for each combination of the four varied parameters: the autonomous element of the REER, the Verdoorn coefficient and the REER elasticity. Table 4 showcases the results: first, as expected, the rate of identified relationships rises with the REER elasticity of exports. Second, also unsurprisingly, permanent export shocks reduce the rate of false positives when there is no true price elasticity and the ARDL’s power when there is: autonomous and persistent shocks break down the cointegrated relationship between exports, the price variable (when it exists) and foreign demand.

Third, a Verdoorn effect increases the rate of false positives, though the effect on the REER is weak enough that the proportion remains small at 10-12% (vs 7-8% without Verdoorn). What is more notable is that the Verdoorn effect increases the power of the ARDL in the case where there is a true price elasticity.

Fourth, and importantly: the DGP2 results lend some support to our hypothesis that the significant relationships identified between ULC-REERs and exports are likely to have more to do with the non-wage share components of the REER, the deflator and the NEER, as those are the components that drive the REER’s movements. The power of the ARDL to detect true price elasticity is quite low when the ULC-REER is only affected by the wage-induced wage share movements but increases dramatically when the autonomous element of the REER is added (see lines 5-6 vs 2-3 in the last column).

To explore this hypothesis further, we set up a third data-generating process, DGP3, identical to DGP2 but for the export equation, which has separate price coefficients for the wage share and non-wage share components of the REER. The elasticities and variances of the two components are calibrated on the sample used throughout the paper, ten euro area countries<sup>3</sup>, 1995-2021. First, the variability of the ULC-REER matches that in the sample, as does the wage share's share of the variability. Second, the elasticities are constructed on the basis of the cost structure of manufacturing exports reconstructed from OECD IO tables: an assumed unit price elasticity of exports yields a -0.37 wage elasticity (proportional to the domestic labour content, which includes labour cost from non-manufacturing inputs), and an elasticity of -0.68 for the deflator-NEER (price) component, proportional to the domestic value-added content of exports.

The export equation of DGP3 is:

$$x_t = \beta_{fd} f d_t + \theta_{ws} w s_t + \theta_z z_t + v_t \quad (22)$$

The results are reported in Table 5, as for DGP2 with 1'000 runs for each combination of the varied parameters: endogeneity of the wage share (labour hoarding + Verdoorn off or on), true price elasticities (none, price only, wage share only, both).

Table 5: Results of method (ii) estimated on DGP3 runs, 1'000 runs per cell

| Case            | Benchmark | ULC-REER      |         | GDP-deflator REER |         |
|-----------------|-----------|---------------|---------|-------------------|---------|
|                 |           | median        | % both  | median            | % both  |
| None            | 0         | 0.00 / -0.19  | 6 / 10  | 0.01 / 0.01       | 6 / 6   |
| Price only      | -0.53     | -0.55 / -0.64 | 25 / 24 | -0.69 / -0.67     | 32 / 31 |
| Wage share only | -0.08     | -0.08 / -0.30 | 6 / 12  | 0.03 / 0.01       | 6 / 6   |
| Both            | -0.61     | -0.63 / -0.74 | 33 / 26 | -0.67 / -0.74     | 31 / 33 |

*Notes:* Each cell shows results for endogeneity off and on ( $\varphi = 1$  and  $\beta_V = 0$ ;  $\varphi = 0.2$  and  $\beta_V = 0.5$ ).  $T = 107$  quarters,  $\rho_v = 0.85$ . The first column indicates which component of the ULC-REER has a true price elasticity ( $\theta_z = -0.68$  for the price component,  $\theta_{ws} = -0.37$  for the wage share). The benchmark column indicates the expected ULC-REER coefficient with no endogeneity,  $(1-s)\theta_z + s\theta_{ws}$ , where  $s = 0.22$  is the wage share's share of the ULC-REER's variance. The next four columns indicate the median estimated elasticities and the % of runs with significant negative coefficients and a passed bounds test, using as regressor respectively the ULC-REER (wage share + GDP-deflator REER) and the GDP-deflator REER.

First, our main hypothesis, that the wage share's independent effect on exports should not translate into a large ULC-REER coefficient, is verified. Because it accounts for 22% of the variance of the REER and realistically only has an elasticity proportional to its share of output prices, its own elasticity is detected by the ARDL in 6% of runs, barely above the nominal rate. With the added endogeneity, power jumps up to 12%, but the estimated ULC-REER coefficient is still quite small at -0.3. With no endogeneity, the price elasticity of -0.68 shows up in a slightly biased ULC-REER coefficient of -0.55 and an unbiased GDP-REER coefficient. As before, wage share endogeneity biases down the ULC-REER coefficient, to -0.19 in the no true price elasticity setting (10% rate of false positives), -0.64 vs. -0.55 in the no endogeneity setting when there is a price elasticity but no wage share elasticity. As a result of the previous two elements, the ULC-REER coefficient is only of significantly lower magnitude than the GDP-REER coefficient in the case with no endogeneity and no true wage share elasticity, suggesting this pattern, sometimes

<sup>3</sup>Excluding Ireland because of its massive profit-shifting-induced wage share swings.

found in the literature (e.g. Keil, 2023), could be accounted for by this mechanism: the ULC-REER functioning as a GDP-REER with extra noise introduced by a wage-share orthogonal to export dynamics. Finally, not reported in the table for reasons of legibility: splitting the ULC-REER into its wage-share and GDP-deflator components and using both as regressors does not solve the problem. When ULCs are endogenous, the bias is concentrated in the wage-share coefficient, which takes large negative values ( $-0.8$  to  $-1.2$ ) whatever the true effect.

## 6 Discussion

The simulation-based analyses in the previous section yield several conclusions of interest for the use of ULCs and ULC-REERs as price variables in export equations, relating to the features identified in section 4: the potential endogeneity and the mixing of a wage share and a price-based REER.

First, the endogeneity problem: it is shown to very much be a problem for methods (i), (iii) and to some extent (iv). In setups with no true price effect, it reliably produces false positives using OLS on differences and panel fixed effects, and can fool the mean group panel ARDL methods depending on its precise lag specification and whether the trimmed mean group estimator is used. The endogeneity always produces bias in the price coefficient, which is not enough, however, to produce reliably spurious results in single-country ARDL with bounds test setups, as it only produces a transitory relationship between exports and the ULC-REER, not a permanent one. Because labour hoarding makes the ULC-REER move as a mirror image of exports within the quarter, the contemporaneous change in the REER absorbs the adjustment of exports towards equilibrium. The error correction term loses explanatory power, and the bounds test rarely rejects the null of no levels relationship. In the panel setting, pooling reduces the standard errors, making the biased price term and the error correction term significant.

These results suggest that ULC-REERs should be used with caution with the methods particularly subject to the mechanical endogeneity of short-term ULC changes to value added dynamics. Long-run levels estimates are less subject to this problem, though they can be vulnerable in panel settings. Even when it does not reliably produce spurious results, the procyclicality of the wage share biases the estimated price coefficients. Analysing other methods used in the literature, such as three-stage least squares or vector error correction models, would further clarify the extent of the problem posed by ULCs' endogeneity to exports.

The second implication of ULC-REER construction for export equation (or value share equation) estimates pertains not to the wage share, but to the GDP-REER component. Because this second component accounts for most of the ULC-REER's variance and the wage-elasticity of exports is likely to be limited by wages' modest share of exporters' costs, the independent effect of the wage share on exports is unlikely to be detected with ULC-REERs as a price variable.

Taken together, these two points raise questions about the pertinence of ULC-REERs as price variables: the wage share's own effect is unlikely to be large enough to be detectable and it may add noise, diluting the effect of the GDP-REER (if the endogeneity effect is small) or introducing bias (if it is not).

There is a further theoretical reason for doubting the pertinence of ULC-REER as a price indicator. In DGP3, we include an independent price elasticity for the wage share. The existence

of such an elasticity is the implicit assumption underlying the use of ULC-REERs: if there is no independent role for the wage share, one should simply use the remainder of the ULC-REER, i.e. a GDP deflator-based REER instead of a whole economy ULC-based one, or a manufacturing value added deflator-based REER instead of the manufacturing ULC-deflated REER.

Yet an independent wage share effect on exports is somewhat implausible and arguably not a price effect: if wages rise, leading to a rise in prices, the wage share should remain stable (or even fall somewhat, given a stable markup) and the rise in prices is to be found in the price deflator, not the wage share. A direct wage share effect on exports would have to go through a profit squeeze mechanism, with rising wages squeezing margins, prices not being raised to preserve market share, and either lower entry into the market or higher exit from it. Wages would then have an effect not on price competitiveness per se, but on return on investment, leading to a relocation of production. This effect would likely be longer-term, given the time needed for investment decisions to materialise, and might be offset by countervailing forces, e.g. wage rises incentivising businesses to invest in productivity enhancing technology given the already existing agglomeration economies in a given activity where it happens to be located. The relative weight of these effects would need to be studied with finer-grained methods.

Finally, our results have some bearing on the price vs non-price competitiveness debate and the issue of measuring non-price competitiveness: in our simulations, the identification of ARDL models is impeded by permanent exogenous (“non-price competitiveness”, if you will) shocks to exports, which break the relationship between exports, foreign demand and (when there is one) price variables. Together with the successful estimation of standard export equations, this suggests that non-price competitiveness factors are very stable and absorbed into the income elasticities, as a permanent NPC shock would break the estimated cointegrated relationships. It likely helps, in that regard, that these models are usually estimated on fairly short samples.

## 7 Conclusion

In this article, we identify a series of potential problems posed by the price variables used in measuring competitiveness, REERs and especially the associated deflators. First, at the root of the issue is the fundamental problem: given diversified economies trading heterogeneous goods and specialising in different stages of complex value chains, the direct comparison of levels of *relevant* relative prices, i.e. relative prices of comparable goods or services, is impossible.

We argue the methods used in the literature implicitly adopt one of two ways out of this impossibility: analysing the effects of changes in relative prices or focusing on a key price, that of labour. The first approach corresponds to the use of price indices, and the second to analyses in terms of unit labour costs.

As we argue in section three, no available price index focuses squarely on the costs relevant to exporting firms and while the indices do largely move together in the longer run, they are often relatively decorrelated in the short run. This may in part account for the inconsistent estimates of price elasticities in the literature (see the discussion in Keil, 2023).

We then discuss three issues with unit labour costs as a price variable: the small and variable proportion they occupy in exporters’ costs, their endogeneity to value added swings and their incorporation of a price index which in fact dominates their variations.

The latter two issues are explored using Monte Carlo simulations with simple models of export dynamics. The results show that most methods tested – OLS in differences, fixed effects models, some specifications of panel ARDL with a mean group estimator – are particularly prone to the endogeneity problem, reliably producing spurious results, while others – ARDL models with bounds tests and some panel ARDL specifications – are not prone to producing spurious results but are nonetheless biased by the regressor’s endogeneity. These results suggest caution is indicated when choosing a methodology for estimating export equations. They need to be supplemented, however, with careful empirical investigation as our DGPs are very stylised and the size of the spurious effect depends on the calibration of the hoarding term and the wage noise.

Finally, both our analysis of empirical ULC-REER dynamics and the simulation results suggest that the price elasticities detected in the literature are likely to result from the components of the ULC-REER that are shared with GDP deflator-based REERs, casting doubt on the rationale behind the use of ULC-deflated REERs. Not only is their wage share component unlikely for theoretical reasons to have a direct effect on exports, but such a hypothetical effect, if of plausible size, would also be unlikely to be detected, and ULC-REERs may introduce bias as a result of their endogeneity to export performance. The testing of various REERs and possibly the wage share as an separate regressor would thus shed more light on the causalities at work and what drives the identified elasticities.

More work is needed, however, to ascertain which methods and which settings are likely to produce spurious results because of this bias. As they stand, our results suggest some caution with the interpretation of significant ULC or ULC-REER elasticities as evidence of a causal link from wages to exports.

In terms of the direct effects of wages and the wage share on export performance, fine-grained analyses of firms’ investment and localisation decisions would shed light on the question.

## References

- Amiti, M., Itskhoki, O., and Konings, J. (2013). Importers, Exporters, and Exchange Rate Disconnect. NBER Working Paper 18615, National Bureau of Economic Research.
- Baccaro, L. and Pontusson, J. (2016). Rethinking Comparative Political Economy: The Growth Model Perspective. *Politics & Society*, 44(2):175–207.
- Baccaro, L. and Tober, T. (2022). The role of wages in the Eurozone. *Review of International Political Economy*, 29(4):1263–1286.
- Balassa, B. (1964). The Purchasing-Power Parity Doctrine: A Reappraisal. *Journal of Political Economy*, 72(6):584–596.
- Baumol, W. J. and Bowen, W. G. (1965). On the Performing Arts: The Anatomy of their Economic Problems. *The American Economic Review*, 55(1/2):43–50.
- Biddle, J. E. (2014). Retrospectives: The Cyclical Behavior of Labor Productivity and the Emergence of the Labor Hoarding Concept. *Journal of Economic Perspectives*, 28(2):197–212.
- Boggio, L. and Barbieri, L. (2017). International competitiveness in post-Keynesian growth theory: controversies and empirical evidence. *Cambridge Journal of Economics*, 41(1):25–47.

- Bottega, A. and Romero, J. P. (2021). Innovation, export performance and trade elasticities across different sectors. *Structural Change and Economic Dynamics*, 58:174–184.
- Dossche, M., Gazzani, A., and Lewis, V. (2023). Labor adjustment and productivity in the OECD. *Review of Economic Dynamics*, 47:111–130.
- Ederer, S. and Reschenhofer, P. (2018). Macroeconomic imbalances and structural change in the EMU. *Structural Change and Economic Dynamics*, 46:59–69.
- Felipe, J. and Kumar, U. (2014). Unit labor costs in the eurozone: the competitiveness debate again. *Review of Keynesian Economics*, 2(4):490–507.
- Giordano, C. and Zollino, F. (2016). Shedding Light on Price- and Non-price-competitiveness Determinants of Foreign Trade in the Four Largest Euro-area Countries. *Review of International Economics*, 24(3):604–634.
- Gräbner, C., Heimberger, P., and Kapeller, J. (2020a). Do the "smart kids" catch up? Technological capabilities, globalisation and economic growth. Technical Report 9, ifso working paper.
- Gräbner, C., Heimberger, P., Kapeller, J., and Schütz, B. (2020b). Is the Eurozone disintegrating? Macroeconomic divergence, structural polarisation, trade and fragility. *Cambridge Journal of Economics*, 44(3):647–669.
- Gräbner-Radkowsch, C., Tamesberger, D., Heimberger, P., Kapelari, T., and Kapeller, J. (2022). Trade models in the European Union. *Economic Annals*, 67(235):7–36.
- Hall, P. A. and Soskice, D., editors (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford University Press.
- Kapeller, J., Gräbner-Radkowsch, C., and Heimberger, P. (2019). Economic polarisation in Europe: Causes and policy options. Technical Report 440, wiiw Research Report.
- Keil, S. (2023). The challenging estimation of trade elasticities: Tackling the inconclusive Eurozone evidence. *The World Economy*, 46(5):1235–1263.
- Keil, S. (2024a). Assessing the role of non-price factors: Shedding new light on the European competitiveness puzzle. *International Economics*, 178(100495):1–21.
- Keil, S. (2024b). Competing for manufacturing value added: How strong is competitive cost pressure on sectoral level? *Structural Change and Economic Dynamics*, 69:197–212.
- Kohler, K. and Stockhammer, E. (2022). Growing differently? Financial cycles, austerity, and competitiveness in growth models since the Global Financial Crisis. *Review of International Political Economy*, 29(4):1314–1341.
- Lapavistas, C. (2012). *Crisis in the Eurozone*. Verso Books, London.
- Pariboni, R. and Paternesi Meloni, W. (2022). The political economy of alternative export-led strategies: Exporting differently? Technical Report 80, FMM Working Paper.
- Paternesi Meloni, W. (2018). Italy’s Price Competitiveness: An Empirical Assessment Through Export Elasticities. *Italian Economic Journal*, 4(3):421–462.
- Paternesi Meloni, W. (2021). The price vs. non-price competitiveness conundrum: a post-Keynesian comparative political economy analysis. Technical Report 2109, Post-Keynesian Economics Society Working Paper.

- Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3):289–326.
- Pesaran, M. H. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.
- Samuelson, P. A. (1964). Theoretical Notes on Trade Problems. *The Review of Economics and Statistics*, 46(2):145.
- Shaikh, A. (2016). *Capitalism: competition, conflict, crises*. Oxford University Press, Oxford New York, NY.
- Storm, S. and Naastepad, C. (2016). Myths, Mix-ups, and Mishandlings: Understanding the Eurozone Crisis. *International Journal of Political Economy*, 45(1):46–71.
- Storm, S. and Naastepad, C. W. M. (2015). Europe’s Hunger Games: Income Distribution, Cost Competitiveness and Crisis. *Cambridge Journal of Economics*, 39(3):959–986.
- Thirlwall, A. P. (1979). The Balance of Payments Constraint as an Explanation of International Growth Rate Differences. *PSL Quarterly Review*, 32(128):45–53.

## Annexes

### A1 DGP1, DGP2 and DGP3: specification and calibration

DGP1 is given by equations (5)–(16). DGP2 replaces trend productivity with the Verdoorn equation (20) and the export equation with (21), in which exports depend on the contemporaneous ULC-REER. DGP3 replaces the export equation with (22), in which the wage share and the exogenous component of the REER have separate elasticities. In DGP2 and DGP3 the export equation creates a simultaneity between exports, value added and measured productivity within the quarter, which we solve analytically: productivity reacts to value added with the coefficient  $\kappa = 1 - \varphi + \varphi\beta_V$ , so that exports are the solution of a linear equation in  $x_t$  given the period's shocks and the lagged state.

The DGPs are quarterly and each simulation discards a burn-in of 100 quarters. Table 6 gives the calibration. The employment adjustment speed is calibrated by indirect inference on annual AMECO data for Keil's countries excluding Ireland, 1995–2021: we simulate DGP1 at a given quarterly  $\varphi$ , aggregate the simulated series to annual averages and choose the value of  $\varphi$  for which partial adjustment and Okun-type regressions on the simulated data match their estimates on the actual data. In DGP3 the standard deviations of the wage and  $z$  shocks are set so that the simulated ULC-REER matches two moments of the quarterly ULC-REER in the sample (European Commission, 37 partners, 1995Q1–2021Q3, excluding Ireland, linear trend removed, median across countries): the standard deviation of its level (0.043) and the wage share's share of its variance (0.22). The two elasticities of DGP3 assume a unit price elasticity of exports and are proportional to the vertically integrated domestic content of manufacturing exports computed from the OECD national input-output tables (mean over 1995–2021, excluding Ireland): 0.37 for labour costs and 0.68 for value added.

Table 6: Calibration of DGP1, DGP2 and DGP3 (quarterly)

| Parameter               | Description                     | Value                     | Source / rationale                  |
|-------------------------|---------------------------------|---------------------------|-------------------------------------|
| $\beta_{fd}$            | Income elasticity of exports    | 1.5                       | Set                                 |
| $\mu_{fd}, \sigma_{fd}$ | Foreign demand drift, shocks    | 0.006, 0.006              | Set (about 2.4% a year)             |
| $\rho_v$                | Persistence of export shocks    | 0.85 (1 in DGP2 variants) | Set                                 |
| $\sigma_v$              | Export shocks                   | 0.020                     | Set                                 |
| $\mu_{dd}, \sigma_{dd}$ | Domestic demand drift, shocks   | 0.004, 0.006              | Set                                 |
| $s_x$                   | Share of exports in value added | 0.35                      | Set                                 |
| $g_a$                   | Trend productivity growth       | 0.003                     | Set (about 1.2% a year)             |
| $\varphi$               | Employment adjustment speed     | 1 / 0.5 / 0.2             | 0.2 calibrated (indirect inference) |
| $\sigma_l$              | Employment shocks               | 0.002                     | Set                                 |
| $\sigma_w$              | Wage shocks                     | 0.006 (DGP3: 0.0075)      | Set (DGP3: calibrated)              |
| $\sigma_z$              | Exogenous REER component        | 0 / 0.015 (DGP3: 0.0155)  | Set (DGP3: calibrated)              |
| $\beta_V$               | Verdoorn coefficient            | 0 / 0.3 / 0.6 (DGP3: 0.5) | Set                                 |
| $\theta$                | REER elasticity (DGP2)          | 0 / -0.3 / -1             | Set                                 |
| $\theta_{ws}, \theta_z$ | Elasticities (DGP3)             | -0.37, -0.68              | OECD input-output tables            |

### A2 DGP1p: specification and calibration

DGP1p is annual and simulates 28 countries and 18 manufacturing sectors, i.e. 504 country-sectors, discarding a burn-in of 50 years. It is given by equation (17) for value added; employment

adjusts to its target with the same partial adjustment mechanism as in DGP1, at the annual speed  $\varphi = 0.3$  (calibrated on annual data as for DGP1), and wages follow a random walk with a drift equal to trend productivity growth plus a country and a country-sector shock. Relative ULCs are measured against the average of the same sector across countries, value-added shares against the sector's European total. Table 7 gives the calibration.

Table 7: Calibration of DGP1p (annual)

| Parameter                  | Description                            | Value                                                                    | Source / rationale                                      |
|----------------------------|----------------------------------------|--------------------------------------------------------------------------|---------------------------------------------------------|
| $N_c \times N_s$           | Countries $\times$ sectors             | $28 \times 18$                                                           | Set                                                     |
| $\mu_D, \sigma_D$          | European sector demand: drift, shocks  | 0.02, 0.04                                                               | Set                                                     |
| $\mu_d, \sigma_d$          | Country domestic demand: drift, shocks | 0.02, 0.02                                                               | Set                                                     |
| $x_{cs}$                   | Export share of demand                 | sector export dependence<br>$\times$ country openness, in $[0.05, 0.95]$ | Keil (2024b, Table 1)<br>openness log-normal, s.d. 0.35 |
| $Y_c$                      | Country size (GDP)                     | log-normal, s.d. 1                                                       | Set                                                     |
| $\rho_v, \sigma_v$         | Unit competitiveness shocks            | 0.5, 0.05                                                                | Set                                                     |
| $g_a$                      | Trend productivity growth              | 0.015                                                                    | Set                                                     |
| $\sigma_l$                 | Employment shocks                      | 0.01                                                                     | Set                                                     |
| $\sigma_{wc}, \sigma_{wu}$ | Wage shocks: country, country-sector   | 0.015, 0.015 ( $\times 1$ or $\times 3$ )                                | Set                                                     |
| $\varphi$                  | Employment adjustment speed            | 1 / 0.3                                                                  | 0.3 calibrated                                          |
| $\theta$                   | Relative ULC elasticity                | 0                                                                        | –                                                       |

### A3 Monte Carlo procedure

Each cell of Tables 2, 4 and 5 is based on 1'000 independent simulations of the DGP, each cell of Table 3 on 500. For each simulation, the estimators of section 5 are applied to the simulated series and the tables report medians (means for method (i) in Table 2) of the estimated coefficients and the shares of runs in which the relevant tests are significant at 5%. The simulations were run in R; the scripts are available on request.