UNIVERSITÄT BERN

Faculty of Business, Economics and Social Sciences

Department of Economics

Technology shocks and hours worked: a cross-country analysis

Jacqueline Thomet, Philipp Wegmüller

18-19

June 2018

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach 8573
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Technology shocks and hours worked: a cross-country analysis*

Jacqueline Thomet** Philipp Wegmueller***

June 30, 2018

Abstract

Using a novel data set, we reassess the evidence for (or against) a key implication of the basic RBC model: that aggregate hours worked respond positively to a positive technology shock. Two novel aspects of the analysis are the scope (14 OECD countries) and the inclusion of data on both labor supply margins to analyze the key margin of adjustment in aggregate hours. We show that the short-run response of aggregate hours to a positive technology shock is remarkably similar across countries, with an impact fall in 13 out of 14 countries. In contrast, the decomposition of the aggregate hours results into intensive and extensive margins shows substantial heterogeneity in the labor market dynamics across OECD countries. For instance, movements in the intensive margin are the dominant channel of adjustment in aggregate hours in 5 out of 14 countries of our sample, including France and Japan.

JEL class: E24, E32

Keywords: Structural VAR, technology shocks, aggregate hours worked, labor supply margins, relative price of investment.

*We wish to thank Luca Benati and Fabrice Collard for guidance from start to finish, and Andrea Raffo and Lee Ohanian for making the data available to us. We also thank Mario Forni, Jordi Galí, Christian Myohl, and participants of the Macro Workshop at the University of Bern and of the 2017 Spring Meeting of Young Economists for valuable comments.

**Department of Economics, University of Bern. Email: `thometj@gmail.com`

***State Secretariat for Economic Affairs, Switzerland. Email: `wegmueller.philipp@gmail.com`

1 Introduction

Fueled by the seminal contributions of Kydland and Prescott (1982) and Long and Plosser (1983), much attention has been devoted to estimating the impact of technology shocks on aggregate hours worked. A focal point of the empirical literature has been to test the key predictions of the basic real business cycle (RBC) model, namely that aggregate hours worked per capita rise after a permanent technology shock and that the correlation of the technology-induced components of aggregate hours and labor productivity is positive. Although the empirical evidence is not clear-cut, there is today a firm presumption that—at least for the US—these predictions are rejected by the data.[1]

For countries other than the US, the empirical evidence stands on a more insecure footing. Data availability is a key difficulty: Until recently, consistent quarterly data on hours worked per employee (the intensive margin of labor supply) has not been available for many countries other than the US. For lack of an alternative, the international empirical evidence on the role of technology for labor supply relies on employment (the extensive margin of labor supply) as a proxy for aggregate hours, including for instance the evidence on G7 and Eurozone countries provided by Galí (1999, 2004) or Dupaigne and Fève (2009). Addressing this lack of data, Ohanian and Raffo (2012) have published a homogeneous quarterly data set for 14 OECD countries on aggregate hours worked and its two components, hours per employee and employment. Their data shows that in many countries (and in contrast to the US), much of the cyclical variation in aggregate hours takes place in average hours per employee. This finding suggests that employment is potentially a poor proxy for aggregate hours and puts much of the existing international evidence on the role of technology shocks for aggregate hours into question.

In this paper, we use the data by Ohanian and Raffo (2012) (and a subsequent update provided by the authors) and analyze how the response of aggregate hours to permanent technology shocks compares across industrialized countries. To shed further light on the dynamics of aggregate hours, we disentangle the response of aggregate hours into its reactions along the intensive and extensive margins of labor supply. We do this separately for each country. This exercise is particularly interesting in our cross-country set-up, as the countries we study feature a wide range of labor market institutions and legislation, which translates into different incentives to adjust labor along these margins. To the best of our knowledge, we are the first to perform this exercise with the homogeneous data by Ohanian and Raffo (2012).

Our results are based on a structural VAR model estimated over the 1982Q3–2013Q4 period. Different specifications are considered. In our baseline VAR, labor productivity enters in first differences and is adjusted for trend breaks. Aggregate hours enters in levels and is quadratically detrended (to remove low- to medium-frequency movements). Besides aggregate hours and labor productivity, the VAR further includes information on consumption-to-output and investment-to-output ratios, interest rates, and inflation. For the decomposition of the aggregate hours response into intensive and extensive margins of labor supply, we append information on employment to the system and use generalized least squares to estimate the VAR with linear constraints, following Lütkepohl (2005). Technology shocks are identified as the shocks which explain the maximum fraction of forecast error variance (maxFEV) of labor productivity at the 10-year-horizon, a statistical approach associated with Uhlig (2003, 2004).

[1]Prominent contributions to this debate include Galí (1999), Christiano, Eichenbaum, and Vigfusson (2004), Francis and Ramey (2005), Basu, Fernald, and Kimball (2006), Fisher (2006), and Fernald (2007).

We also extend our empirical analysis and disentangle neutral (N) and investment-specific (I) technological change.[2] This analysis is based on a subsample of five countries for which we are able to collect data of sufficient quality to construct a measure of the relative price of investment (RPI). Following Fisher (2006), the simplifying assumption that technology shocks affect the production of all goods homogeneously may be problematic if N- and I-shocks have permanent effects on labor productivity, but different effects on aggregate hours. Against this background, the extension is a useful robustness exercise. In addition, we also investigate if the relative importance of the two labor supply margins for the aggregate hours response varies depending on the type of technology shock.

Overall, our results confirm across a broad set of countries Galí's (1999) findings. On impact, we find a fall in aggregate hours worked in response to a positive technology shock in 13 out of 14 countries of our sample (Australia is the single exception). The technology-driven components of aggregate hours and labor productivity are negatively correlated in 12 countries. There is more variation in our findings when we decompose our results into intensive and extensive margins of labor supply. Employment is the dominant channel of adjustment in 9 out of 14 countries, including the Anglo-Saxon economies. The opposite holds for Austria, France, Japan, Korea, and Norway: In these economies, most of the aggregate hours response to technology shocks is through movements in average hours per employee. More generally, our decomposition shows that the intensive margin is important for the short-run adjustment of aggregate hours. As the response of employment is typically slow and builds up over time, employment is overall an inadequate proxy for the short-run response of aggregate hours. We also investigate the robustness of our main findings. We show that accounting for low- to medium-frequency movements in hours and labor productivity is crucial to avoid substantial estimation bias when hours enter the VAR in levels, reemphasizing the same point by Fernald (2007), Canova, Lopez-Salido, and Michelacci (2010), or Chaudourne, Fève, and Guay (2014) across a broader set of countries. With respect to the decomposition into neutral and investment-specific technical change, we find that the inclusion of investment-specific technical change has little effect on the qualitative features uncovered under the one-technology assumption.

The remainder is organized as follows. Section 2 describes the data and the empirical methodology applied. Section 3 reports our evidence on the role and importance of permanent technology shocks for the 14 OECD countries of our sample. The robustness of these results is evaluated in Section 3.3. Section 4 reports our results for disentangling neutral and investment-specific technical change for a subsample of 5 countries.

2 Data and methodology

This section outlines our empirical framework. We first introduce the reduced-form VAR. Then, Section 2.2 introduces the data and motivates the data pre-treatment. Section 2.3 outlines how the aggregate hours results are decomposed into movements in the intensive and extensive labor supply margins. The last section describes identification.

[2]This extension follows a long line of work including Greenwood, Hercowitz, and Huffman (1988), Gordon (1990), Greenwood, Hercowitz, and Krusell (2000), Fisher (1999, 2006), Smets and Wouters (2007), or Justiniano, Primiceri, and Tambalotti (2011).

2.1 VAR specification

For each country, we estimate the following VAR of order 4 on quarterly data:

$$Y_t = B_0 + B_1 Y_{t-1} + B_2 Y_{t-2} + B_3 Y_{t-3} + B_4 Y_{t-4} + u_t, \quad with \;\; u_t \sim iid(0, \Omega), \tag{2.1}$$

where $Y_t = [Y_{1t}, ..., Y_{nt}]'$ is a vector of n series observed at time t, B_0 is an $n \times 1$ vector of intercepts, the remaining B's are $n \times n$ coefficient matrices, and u_t is an $n \times 1$ vector of zero-mean innovations with covariance matrix $\mathbb{E}[u_t u_t'] = \Omega$.

Our vector Y_t consists of seven series: labor productivity (lp_t), aggregate hours expressed in per capita terms (H_t), the consumption-to-output ratio (c_t), the investment-to-output ratio (i_t), 3-month interest rates (i_t), inflation (π_t), and employment per capita (n_t). The first two series—labor productivity and aggregate hours—correspond to the minimum-system necessary to identify the effect of technology shocks on aggregate hours. We add four variables to this system to mitigate against omitted variable bias. In particular, we include information on nominal interest rates and inflation to help control for the monetary policy setting, which in sticky price models matters for the transmission of technology shocks (see for instance Galí, López-Salido, and Vallés (2003)). We also include real consumption- and investment-to-output ratios. In general equilibrium models, these ratios are typically jointly determined with labor productivity and hours. For US data, Christiano, Eichenbaum, and Vigfusson (2003) show that the omission of consumption- and investment-to-output ratios can cause specification error sufficiently large to qualitatively affect the inference about the effect of technology shocks. Finally, information on employment is added in the last position of the VAR to allow identifying the key margin of adjustment in aggregate hours. Importantly, this last variable is only included to decompose the aggregate hours response into its two margins and is treated in a special way. Section 2.3 provides further details.

2.2 Data

The baseline data set covers quarterly series for 14 OECD countries over the 1982Q3-2013Q4 period. For some countries of our sample, data is available earlier than 1982Q3. We exclude it because the early 1980s are associated, in many industrialized countries, with important changes in the conduct of monetary policy and the macroeconomic environment more broadly.[3] These changes matter for the propagation of shocks and hence the identified technology shocks.

Data on aggregate hours and its two components—average hours per employee and employment—is obtained from Ohanian and Raffo (2012) and a subsequent update provided by the authors. Their measures combine information from various sources, including national statistical offices, establishment surveys, and household surveys. To ensure consistency, the series are adjusted for differences across countries such as paid vacation or sick days. Basic descriptive statistics for these series are summarized in Appendix Section A.2. Data on population aged 16 to 64, consumption, investment, and output is drawn from the OECD Economic Outlook. Consumption-to-output and investment-to-output ratios correspond to the shares of real private consumption and real gross fixed capital formation (GFCF) of real GDP. We construct a measure of labor productivity based

[3] See for instance Bernanke and Blinder (1992), Sims (1992), Taylor (1993) or Ireland (2000), who present evidence that the Federal Reserve abandoned the money supply rule by 1982Q3 for interest rate targeting.

on information on aggregate hours, real output, and population aged 16 to 64, namely (in logs):

$$lp_t \equiv y_t - (Pop_t + H_t) = y_t - (Pop_t + h_t + n_t) \ .$$

We use the OECD Main Economic Indicators' measure of 3-month nominal rates. For Germany, Ireland, and Korea, this series only starts in the early 1990s and 3-month money market rates are used to expand the sample. Finally, we obtain inflation as $\pi_t = \Delta \log(P_t)$, where P_t corresponds to quarterly seasonally adjusted GDP deflators (P_t) obtained from national statistical offices.

Data pre-treatment is a key issue in any structural VAR analysis. We are particularly concerned with low- to medium-frequency movements in aggregate hours. As described for instance by Francis and Ramey (2006, 2009), low-frequency movements in aggregate hours arise from demographic changes, sectoral shifts, or trends in labor markets such as female participation rates, which need not be related to business-cycle movements. There is an open debate on whether these trends need to be removed. On the one hand, over-differencing may distort the sign of the impact response of aggregate hours to technology shocks (e.g. Christiano et al. (2004)). On the other hand, if low-frequency comovements in the data are not removed, they can completely dominate high-frequency results (such as IRFs), as pointed out by Blanchard and Quah (1989), Fernald (2007), Canova et al. (2010), or Chaudourne et al. (2014).[4] Testing for stationarity does not resolve the matter. As aggregate hours are very persistent, the assumption of stationarity is usually rejected by unit root tests in small samples (see Appendix Table 12). Yet, from a theoretical viewpoint, one can easily argue that aggregate hours per capita are by definition level-stationary, as one cannot work more than 24 hours a day. In our baseline specification, we follow Fernald (2007) and Canova et al. (2010) and quadratically detrend aggregate hours to allow for intercept heterogeneity. Given the importance of this choice, various robustness exercises are discussed (Section 3.3.1). For instance, we also show results for aggregate hours in levels and for alternative methods of detrending.

A related issue is trend breaks in labor productivity growth. We apply the endogenous break tests of Bai and Perron (1998, 2003) to determine potential breaks in the mean of quarterly labor productivity growth. Results are reported in Appendix Table 15. We find 6 trend breaks: 1991Q2 for Japan, 1991Q3 for Korea, 1995Q4 for Italy, 2003Q1 for France, 2005Q4 for Norway, and 2008Q1 for Finland. Again, we also investigate the robustness of this choice (Section 3.3.2). Finally, for Germany, we control for the trend shift due to the unification in 1991.

2.3 Decomposition

To decompose the aggregate hours results into movements in the intensive and extensive margins of labor supply, we append information on either the extensive or intensive margin to the system (in our baseline case we append $\hat{n}_t$). Since we do not want this addition to affect the estimation results, we impose the linear restriction that information on $\hat{n}_t$ is not allowed to have a contemporaneous or lagged impact on all other variables of the system. With $\hat{n}_t$ ordered last in the VAR, this restriction amounts to imposing zeros in positions $B_j(1 : n-1, n) \,\forall\, B_j, j > 0$. The VAR model subject to linear constraints is estimated with *estimated* generalized least squares (EGLS) following Lütkepohl

[4]This clearly matters: It is a widely discussed feature of US data that for many specifications, the estimated response of aggregate hours to permanent technology shocks changes its sign depending on how aggregate hours enter the VAR (see e.g. Christiano et al. (2003)). Fernald (2007) and Canova et al. (2010) explain the sign-reversal by level shifts in the data. In the US, high aggregate hours growth in the 1990s coincides with high labor productivity growth. If the level shifts are not accounted for, the positive comovement is partly identified as permanent technology shocks.

(2005). Further details can be found in Appendix Section A.1.1.

To sum up, our baseline system has $Y_t = [\Delta lp_t, \hat{H}_t, cy_t, iy_t, i_t, \pi_t, \hat{n}_t]'$, where $\Delta \equiv 1 - L$ is the first difference operator and a hat denotes quadratically detrended variables. All variables except i_t are measured in logs. In the remainder, to highlight that the last variable in the VAR is treated in a special way, we refer to this system as "*6+1-specification*".

2.4 Identification

After estimating the system in (2.1), we translate the vector of prediction errors of the reduced-form VAR (u_t) into a vector of economically meaningful orthogonal structural shocks (ϵ_t),

$$u_t = A\epsilon_t, \quad with \quad \mathbb{E}_t[\epsilon_t \epsilon_t'] = I. \tag{2.2}$$

From equations (2.1) and (2.2) follows:

$$\Omega = \mathbb{E}[u_t u_t'] = A\mathbb{E}[\epsilon_t \epsilon_t']A' = AA'. \tag{2.3}$$

To identify the coefficients in A, theoretical restrictions must be imposed to reduce the number of unknown structural parameters to be less than or equal to the number of estimated parameters of Ω. For our purposes, we will only identify the technology shock—either under a one-technology assumption (this section) or later distinguishing between embodied and disembodied technological change (Section 4). No additional assumptions are made to separately identify the remainder—the "non-technology shocks". In our baseline case, identification is achieved via the maximum fraction of forecast error variance (maxFEV) approach, pioneered by Uhlig (2003, 2004).[5] The idea of the maxFEV approach is to search for innovations that explain the maximum amount of FEV of a specified variable either at a target horizon or over a specified forecast horizon. In our baseline case, we identify technology shocks as the shocks which explain the maximum fraction of forecast error variance of labor productivity at 40 quarters. Appendix Section A.1.2 provides further information on the methodology.

3 Evidence

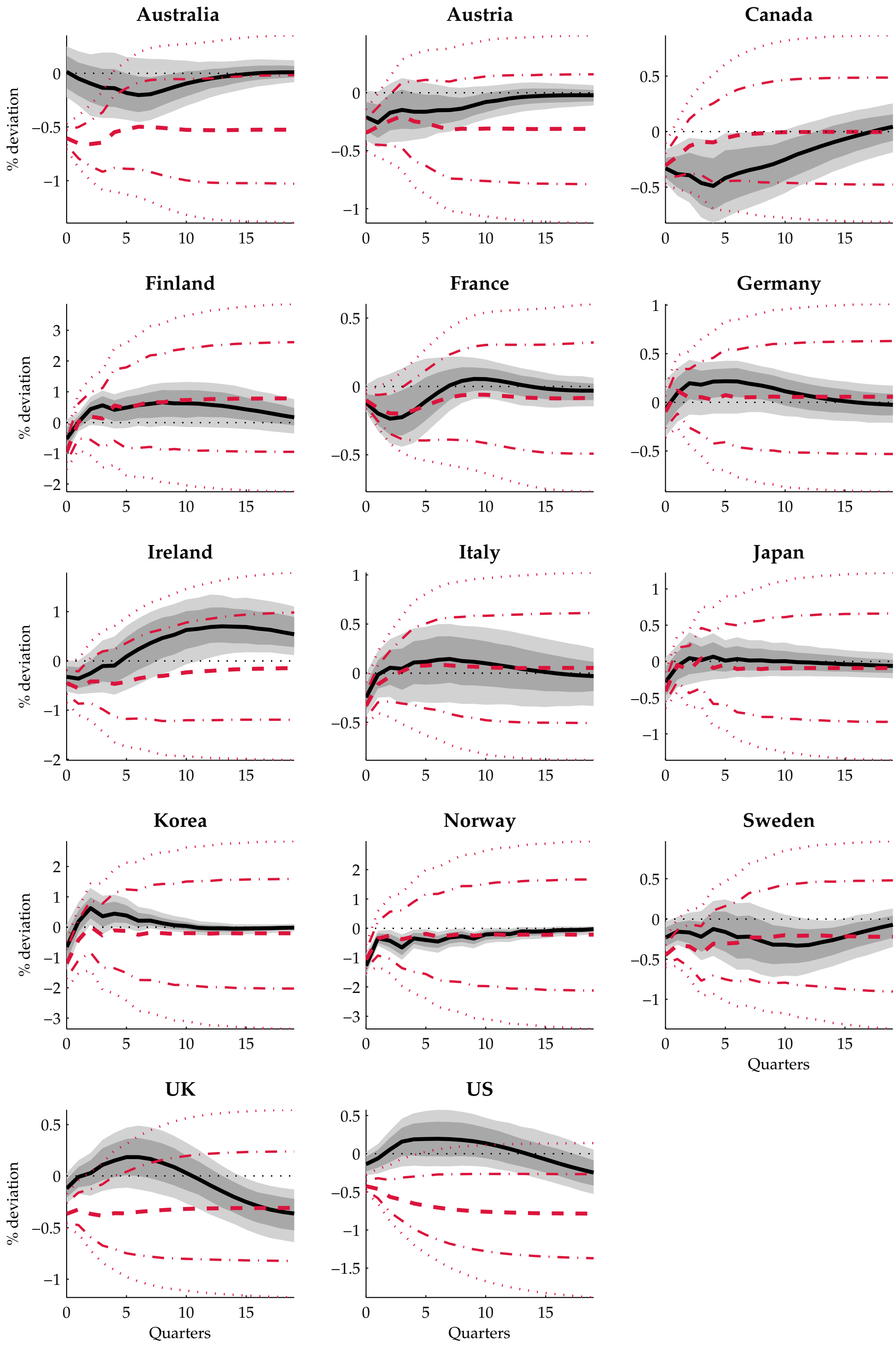
This section presents our cross-country evidence on the labor market response to permanent technology shocks. First, Section 3.1 shows IRFs and FEV decompositions for aggregate hours. Then, Section 3.2 discusses the decomposition of results into adjustments along intensive and extensive margins of labor supply. Section 3.3 evaluates the robustness of our findings.

3.1 Aggregate hours

Figure 1 depicts the median dynamic response of aggregate hours to a one-standard-deviation positive technology shock together with 5th, 16th, 84th, and 95th percentiles. Solid black lines with shaded bands correspond to our 6+1-baseline specification. For comparison purposes, the figure

[5] We choose the maxFEV over long-run restrictions (e.g. applied in Galí (1999)) since it does not suffer from small sample bias in computing impulse responses (e.g. Faust and Leeper (1997)). In particular, Francis, Owyang, Roush, and DiCecio (2014) show with Monte Carlo experiments that the maxFEV identification outperforms identification via long-run restrictions as it reduces the bias in short-run IRFs and raises estimation precision.

Figure 1: *Dynamic response of aggregate hours to a positive technology shock*



Note: Solid black lines with shaded bands depict results for the 6+1-variate baseline specification (*i.e.* quadratically detrended aggregate hours, labor productivity growth adjusted for breaks, identification *via* maxFEV). Dashed red lines depict results for 2+1-specification (*i.e.* aggregate hours and labor productivity in first differences, identification *via* long-run restrictions). Confidence bands are obtained with 1000 bootstrap replications and cover 68% and 90%. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

also depicts results in the spirit of Galí (1999) (red dashed and dotted lines).[6] We view these latter results as a useful historical benchmark to compare our results to.

For the 6+1-specification, the median impact response of aggregate hours is negative in 13 out of 14 countries, and significantly so (at the 68% level) for 12 countries. The single exception is Australia, where the impact response is positive but insignificant. In most countries, the negative response is short-lived. In many instances, aggregate hours revert back to zero within 10 quarters, sometimes after turning positive after the initial fall (e.g. Germany or Korea). At our cut-off point of 20 quarters, hours remain significantly negative in the case of the UK and the US. For the UK, hours revert slowly back to zero; for the US, they remain negative but small. As we discuss in more detail in Section 3.3, the negative impact responses hold for a number of alternative specifications considered.

When we compare our baseline results to the 2+1-specification in the spirit of Galí (1999), three findings stand out. *First*, impact responses are similar across both specifications. For the 2+1-specification, the impact response of aggregate hours is negative in all 14 countries considered. For 13 countries, the fall in hours is significant at the 68%-level. *Second*, a notable difference across the two specifications is the reduction in econometric uncertainty for the 6+1-specification. This reduction in econometric uncertainty is almost entirely due to entering aggregate hours in levels (6+1-specification) rather than in first-differences (2+1-specification). The change in the identification method plays only a minor role.[7] *Third*, the figure depicts almost the same dynamics in the median responses across both specifications, except for Austria and the Anglo-Saxon economies (Australia, Canada, the UK, and the US). In particular, for Australia, Austria, the UK, and the US, we obtain persistent negative effects of positive permanent technology shocks on aggregate hours in the 2+1-specification (significant at the 68% level for Australia and the US). While not shown, the negative effects also continue to hold after our cut-off point of 20 quarters. Except for the US, the persistent negative impact response disappears when including either information on interest rates and inflation or investment- and consumption-to-output ratios (or both, as in our baseline specification), which hints at omitted variable bias in our 2+1-specification and motivates to look at richer specifications.

Another way of looking at the cross-country evidence presented in the IRFs is to compare unconditional and conditional correlation estimates between the cyclical components of labor productivity and aggregate hours. Table 1 summarizes our findings. The table shows that in the data, the unconditional correlation between aggregate hours and labor productivity growth is significantly negative in 10 out of the 14 countries of our sample. Only Germany displays a positive but non-significant correlation. In terms of the technology-driven components of aggregate hours and labor productivity, in our baseline, the conditional correlations are significantly negative for 10 out of 14 countries.[8] In line with the evidence presented in Figure 1 (IRFs), for all countries, the size

[6]These results are reported due to their importance in the literature. We refer to the system as 2+1-specification, as information on n_t is appended to decompose results into the intensive and extensive margins of labor supply. More specifically, we have $Y_t = [\Delta lp_t, \Delta H_t, \Delta n_t]$. As in Galí (1999), technology shocks are identified based on long-run restrictions. Appendix Section A.1.3 provides further details.

[7]To analyze the reason for the reduction in economic uncertainty, we have estimated the 2+1- and 6+1-specifications for both identification schemes. In the 2+1-specification, results are almost identical irrespective of whether shocks are identified based on the maxFEV approach or based on long-run restrictions (not plotted). For the 6+1-specification, results are similar across identification schemes, except Australia (see Appendix Figure 12).

[8]This type of evidence has been used by Galí (1999) to discriminate among RBC and New Keynesian models, as standard RBC models imply a positive comovement between the technology-driven components of aggregate hours and labor productivity.

Table 1: *Conditional and unconditional correlations for aggregate hours*

	6+1-specification (baseline)						**2+1-specification**					
			Conditional on:						Conditional on:			
Country	Uncond.		Residual		Technology		Uncond.		Residual		Technology	
AU	-0.61	**	-0.68	**	-0.57	**	-0.61	**	0.07		-0.89	**
AT	-0.28	**	-0.23	**	-0.78	**	-0.28	**	-0.07		-0.89	**
CA	-0.14		0.37	++	-0.86	**	-0.14		0.33	++	-0.46	**
FI	-0.33	**	-0.43	**	-0.11		-0.28	**	-0.54	**	-0.48	**
FR	-0.13		0.00		-0.68	**	-0.20	**	-0.35	**	-0.73	**
DE	0.08		-0.10		0.34	++	0.08		-0.60	**	-0.06	
IE	-0.31	**	-0.11		-0.29	**	-0.31	**	0.53	++	-0.83	**
IT	-0.19	**	-0.15	*	-0.27	**	-0.24	**	0.12		-0.36	**
JP	-0.17	**	-0.11		-0.63	**	-0.12		0.05		-0.66	**
KR	-0.56	**	-0.73	**	-0.33	**	-0.57	**	-0.38	**	-0.87	**
NO	-0.72	**	-0.48	**	-0.95	**	-0.72	**	-0.60	**	-0.90	**
SE	-0.11		0.12		-0.42	**	-0.11		0.51	++	-0.85	**
UK	-0.30	**	-0.56	**	-0.01		-0.30	**	0.21	++	-0.97	**
US	-0.57	**	-0.76	**	0.06		-0.57	**	0.02		-0.88	**

Note: **/++ indicates significance at 95%-level; */+ indicates significance at 68%-level (based on 1000 bootstrap replications). Moments are based on cyclical components extracted via HP-filter with $\lambda = 1600$. Table 5 reports the decomposition of the results in terms of intensive and extensive labor supply margins.

of the negative correlation is reduced in the baseline specification compared to the 2+1-specification. Results across the two specifications vary even more strongly in terms of the conditional correlations based on the non-technology component (the residual component). This is not unexpected, since the information from the augmentation of variables primarily affects this residual component. In our baseline specification, we find a significant negative correlation between the non-technology components of aggregate hours and labor productivity in 8 out of 14 countries.

Table 2 summarizes median estimates for the fraction of FEV of aggregate hours and labor productivity explained by permanent technology shocks. The table shows estimates at horizons 0 (the impact), 4, 8, 16, and 32 quarters. Overall, the table shows substantial heterogeneity in the role of technology shocks in accounting for fluctuations in aggregate hours across countries. Focusing first on aggregate hours in our baseline specification, on impact, results range from a low of 4% (Australia and Germany) to a high of 48% (Norway). On average, the fraction of FEV of aggregate hours explained by technology shocks remains roughly constant over time, increasing from an average of 15% on impact to 17% after 8 years. A few exceptions aside, the fraction of FEV explained by technology shocks is higher for labor productivity. On impact, results range from a low of 12% (Austria) to a high of 77% (Norway). At the 2-8 year horizon, the fraction of FEV of labor productivity explained by technology shocks lies (well)-above 30% in 9 countries of our sample. The lowest numbers are found for Australia, Austria, and Finland and to a lesser extent Germany and France. As to the comparison of results between the 6+1- and 2+1-specifications, the table shows that going to a larger scale system considerably reduces the role played by technology shocks, both for aggregate hours as well as labor productivity.

To sum up, we find that the impact response of aggregate hours to positive, permanent technology shocks is similar across countries, with an impact fall in 13 out of 14 countries of our sample. In most instances, we find a negative comovement between the technology-induced components of aggregate hours and labor productivity. At the same time, there is substantial cross-country heterogeneity in the quantitative importance of technology shocks for fluctuations in aggregate hours and labor productivity.

Table 2: *Fraction of FEV of aggregate hours and labor productivity explained by technology shocks (in percent)*

	Hours worked										**Labor productivity**									
	Baseline					**2+1-specification**					**Baseline**					**2+1-specification**				
Country	**0**	**4**	**8**	**16**	**32**	**0**	**4**	**8**	**16**	**32**	**0**	**4**	**8**	**16**	**32**	**0**	**4**	**8**	**16**	**32**
AU	4	7	11	12	12	76	67	66	66	66	19	17	17	17	17	89	83	83	82	82
AT	9	10	12	13	14	33	32	32	33	33	12	17	17	17	17	96	87	87	87	87
CA	34	27	26	25	26	30	29	29	29	29	59	48	45	44	44	98	91	90	90	90
FI	15	15	14	15	15	31	45	45	45	45	27	18	17	17	17	90	79	78	78	78
FR	21	16	15	16	16	26	15	18	18	18	33	28	26	25	25	92	89	88	88	88
DE	4	13	18	20	23	3	13	14	14	14	19	19	19	19	20	95	89	88	88	88
IE	14	10	11	18	19	21	22	22	23	23	57	48	45	44	44	95	90	89	89	89
IT	14	8	8	8	9	25	33	33	33	33	43	34	33	32	32	96	90	90	90	90
JP	9	8	8	9	12	22	37	37	37	37	46	32	30	30	30	98	88	88	88	88
KR	10	16	16	15	15	37	46	46	46	46	52	38	37	37	37	96	90	89	89	89
NO	48	35	31	30	28	45	39	38	38	38	77	49	46	46	46	97	80	77	77	77
SE	17	11	10	15	16	48	43	43	43	43	42	34	32	31	31	91	85	82	82	81
UK	8	6	7	12	20	54	41	40	40	40	64	52	50	49	49	95	87	86	86	86
US	8	7	7	11	25	77	49	47	47	47	56	41	40	39	39	61	57	57	57	57

3.2 Intensive and extensive margins of labor supply

We now turn to the decomposition of our aggregate hours results into adjustments in intensive and extensive margins of labor supply. As a motivating observation, Table 3 reports (in the first 3 rows) the percentage of the variability in the cyclical component of aggregate hours over the 1982Q3–2013Q4 period that is accounted for by either movement in employment n_t, hours per employee h_t, or their covariance. This decomposition follows an exercise by Hansen (1985) for the US. In line with Hansen's results, the table shows that for the US, over 60% of the variance of aggregate hours can be attributed to adjustments in employment per capita. Only 10% is due to adjustments in hours per employee. The numbers are similar for the other Anglo-Saxon countries. Interestingly, the pattern is not the same for Austria, France, Japan, Korea, and Norway, where fluctuations in hours per employee account for a higher fraction of the overall variance in aggregate hours than fluctuations in employment. Against this background, a natural question to ask is to what extent the two margins of labor can account for the aggregate hours response to technology shocks described in the last section.

A few exceptions aside, we find that on impact of the technology shock, most of the adjustment in aggregate hours takes place along the intensive margin. Further, the importance of the extensive margin picks up over time. Overall, the extensive margin plays the dominant role in explaining aggregate hours adjustments to technology shocks in all countries but Austria, France, Japan, and Norway. The first finding can be seen in Table 4, which summarizes our decomposition of the median

Table 3: *Variability in aggregate hours due to fluctuations in the two margins*

	Mean	AU	AT	CA	FI	FR	DE	IE	IT	JP	KR	NO	SE	UK	US
Unconditional															
n_t	54	61	28	53	63	50	69	71	60	24	33	38	94	55	61
h_t	36	24	60	17	36	51	54	15	25	61	45	56	30	14	10
$Cov(h_t, n_t)$	10	15	12	30	1	-0	-23	13	14	15	21	6	-24	30	29
Conditional on technology shock															
n_t	51	71	14	39	62	31	99	92	37	18	39	4	86	72	53
h_t	41	9	64	18	28	56	63	16	24	102	46	80	46	9	15
$Cov(h_t, n_t)$	7	20	21	43	10	14	-58	-9	39	-23	14	16	-34	19	34

Note: Numbers are based on the variance decomposition (in logs): $\text{Var}(H_t) = \text{Var}(h_t) + \text{Var}(n_t) + 2\text{Cov}(h_t, n_t)$. Cyclical components are extracted based on the HP-filter with $\lambda = 1600$.

Table 4: *Median impact response of employment and hours per employee to a positive technology shock*

	6+1-specification (baseline)				2+1-specification			
Country	n_t		h_t		n_t		h_t	
AU	-0.02		0.04		-0.22	**	-0.38	**
AT	-0.01		-0.20	**	-0.02		-0.32	**
CA	-0.13	**	-0.20	**	-0.09	*	-0.22	**
FI	-0.12	*	-0.41	*	-0.07		-0.90	**
FR	-0.01		-0.11	**	0.00		-0.10	**
DE	0.03		-0.10		0.05	+	-0.14	*
IE	-0.19	*	-0.13	**	-0.29	**	-0.16	**
IT	-0.13	**	-0.12	*	-0.12	**	-0.21	**
JP	-0.07	*	-0.22	*	-0.06	**	-0.35	**
KR	0.07		-0.72	**	-0.05		-1.15	**
NO	-0.12	**	-1.16	**	-0.09	**	-0.98	**
SE	-0.03		-0.20	**	-0.07	*	-0.38	**
UK	-0.01		-0.11	**	-0.09	**	-0.27	**
US	-0.06	*	-0.08	*	-0.21	**	-0.22	**

Note: **/++ indicates significance at 95%-level, */+ indicates significance at 68%-level (based on 1000 bootstrap replications).

impact response of aggregate hours reported in Figure 1 for the 6+1-specification. The table shows that the intensive margin response is higher (in absolute terms) than the extensive margin response in all countries except Ireland and Italy. Stated differently, we find that in the very short-run, the fall in aggregate hours depicted in Figure 1 (present for all countries except AU) is mainly driven by a decrease in hours per employee. As to the finding of an increase in the importance of extensive margin adjustments over time, consider for instance Appendix Figures 9–11, which show the analog results to Figures 1 (IRFs) and 8 (FEV decompositions) in terms of employment and hours per employee. Appendix Figures 9 and 10 (the IRFs for the two margins) show that the response of employment to technology shocks tends to build up slowly over time, while the response of hours per employee reverts back to zero rather quickly. Alternatively, Figure 11 shows that for all countries, the fraction of FEV of employment accounted for by technology shocks is increasing over time.[9]

We provide two alternative measures of the overall importance of the two labor supply margins for the aggregate hours response. Both measures are based on counterfactual simulations in which we extract the technology-driven components of the series. First, rows 4-6 in Table 3 repeat the variance decomposition of aggregate hours for the technology-driven components. Except for Austria, France, Japan, Korea, and Norway, we find that variation in employment accounts for the bulk of the adjustments in aggregate hours. Movements in hours per employee appear least important in Australia, the UK, and the US. Another interesting finding is that the technology-driven cyclical components of intensive and extensive margin adjustments are negatively correlated in Germany, Ireland, Japan, and Sweden; and positive in all remaining countries. As a second measure, Table 5 summarizes unconditional and conditional correlation estimates between labor productivity and each margin of labor supply. The table shows that the negative unconditional correlation between aggregate hours and labor productivity (as summarized in Table 1) also broadly holds across the two components of aggregate hours individually. In only two cases we observe a change in sign, namely for employment in France and hours per employee in Sweden. In 10 out of 14 countries, the

[9] The decomposition of the IRFs in Appendix Figures 9 and 10 allows to further qualify our finding that the median responses across the 2+1-specification and 6+1-specification are similar in all cases but Austria and the Anglo-Saxon economies. The decomposition of IRFs shows that for Australia, Austria, and the US, the additional information from the augmentation of variables affects both intensive and extensive margins. For Canada and the UK, the difference mainly stems from additional information affecting employment.

Table 5: *Conditional and unconditional correlations for employment and hours per employee*

	Employment per capita						**Hours per employee**					
			Conditional on:						Conditional on:			
Country	Uncond.		Residual		Technology		Uncond.		Residual		Technology	
AU	-0.57	**	-0.63	**	-0.64	**	-0.35	**	-0.40	**	-0.08	
AT	0.28	++	0.30	++	-0.55	**	-0.55	**	-0.53	**	-0.71	**
CA	-0.10		0.32	++	-0.82	**	-0.17	*	0.31	++	-0.81	**
FI	-0.05		0.00		0.04		-0.50	**	-0.63	**	-0.27	**
FR	0.27	++	0.20	++	0.00		-0.45	**	-0.21	**	-0.91	**
DE	0.14		-0.08		0.17	++	-0.06		-0.05		0.22	++
IE	-0.25	**	-0.16	*	0.04		-0.26		0.06		-0.82	**
IT	-0.07		-0.05		-0.23	**	-0.28		-0.22	**	-0.26	**
JP	-0.12		0.02		-0.34	**	-0.15	*	-0.16	*	-0.47	**
KR	-0.18	**	-0.47	**	0.25	++	-0.67	**	-0.68	**	-0.71	**
NO	-0.30	**	-0.18	**	-0.42	**	-0.70	**	-0.44	**	-0.96	**
SE	-0.12		0.06		0.04		0.02		0.10		-0.67	**
UK	-0.31	**	-0.59	**	0.03		-0.18	**	-0.32	**	-0.11	
US	-0.62	**	-0.81	**	0.02		-0.28	**	-0.41	**	0.11	

Note: **/++ indicates significance at 95%-level, */+ indicates significance at 68%-level (based on 1000 bootstrap replications). Moments are based on cyclical components extracted via HP-filter with $\lambda = 1600$. Results correspond to 6+1-variate baseline specification.

unconditional correlation of labor productivity is higher (in absolute terms) with hours per employee than employment. Almost the same finding is obtained when looking at correlation estimates conditional on technology shocks: Again, almost all correlation coefficients are significantly negative, and more so in the intensive than in the extensive margin in 11 out of 14 countries.

Our results highlight interesting differences between countries generally associated with high and low labor market rigidities, respectively. A few exceptions aside, countries in which the intensive margin is important in explaining aggregate hours fluctuations (as highlighted in Table 3) coincide with countries with high firing frictions, as summarized for instance by the OECD ranking of protection against dismissal (printed in Appendix Table 14).[10] Consider for instance the evolution of the FEV of employment between France/Germany (high firing restrictions) and Canada/the US (low firing restrictions), depicted in Appendix Figure 11. For the former two, the fraction of FEV starts near zero on impact and then rises rapidly. In contrast, for the US and Canada, the fraction of FEV is already significantly positive on impact. Stated in terms of the impact response of aggregate hours, the extensive margin accounts for roughly 40% of the impact response in aggregate hours for both Canada and the US. It plays almost no role for both Germany and France (see Table 4). All of these observations are consistent with a higher flexibility of adjusting employment in Canada and the US compared to Germany or France. However, we also want to emphasize that the evidence we present in this regard is not perfect. For instance, Italy ranks similar to France and Germany on the OECD employment protection rankings but has important employment fluctuations.

We want to highlight the following general implications of our decomposition results. First, the fact that most short-run adjustment in aggregate hours takes place along the intensive margin highlights that working with employment as a proxy for aggregate hours can be misleading. For instance, in his analysis for Japan over the 1962Q1–1994Q4 period, Galí (1999) was unable to recover a negative impact response of employment to a positive technology shock. We reproduce

[10]This empirical finding is further explored in Chapter 2 of this thesis. More specifically, the chapter explores the interaction of firing costs and average hours per employee in a search and matching model that features two distinct labor supply choices.

Galí's result with our data set, using his time period and specification. We confirm that the employment response is non-negative—it is near zero and then turns positive. Yet, as expected from our results in Table 4, the impact response turns negative when including the additional information form the intensive margin. Overall, our decomposition results show that the information from the intensive margin are especially important when focusing on the short-run response of aggregate hours. Over longer horizons, employment becomes a better proxy. A second observation is that the observed delay in the reaction in the extensive margin to a positive technology shock appears consistent with the idea that (un)employment evolves as a stock, while hours evolve as a flow. The results are consistent with the idea of frictional labor markets as introduced by Mortensen and Pissarides (1994) and used by Merz (1995) or Andolfatto (1996) (among others) for studying economic fluctuations. In a nutshell, the idea of this approach is that workers take time to find a job and that it is costly for firms to find suitable workers, both generating a delay in the reaction of the extensive margin to technology shocks.

3.3 Robustness

In the following, we consider alternative estimation and specification choices. We first review the sensitivity of our baseline results with respect to the (auxiliary) choices of the number of lags included in the VAR and the horizon of identification. We then focus on data pre-treatment.

3.3.1 Sensitivity of results to auxiliary choices

How important is the horizon of identification, the VAR lag length, and the deflator for finding a negative impact response of aggregate hours to positive technology shocks? Table 6 summarizes the size and significance of the median impact response of aggregate hours by—starting from our baseline specification—altering either the number of lags included in the VAR or the horizon of identification.[11] In particular, the first block reports results for our baseline specification, with a lag

Table 6: *Robustness of baseline SVAR results to lag lengths and forecast horizon*

	(1) Baseline	(2) Lag order				(3) Forecast horizon		
		1	2	8	12	20	60	80
AU	0.01	-0.05	0.02	-0.01	-0.01	-0.08	0.04	0.04
AT	-0.21 *	-0.33 **	-0.27 **	-0.28 *	-0.27 **	-0.26 **	-0.14	-0.07
CA	-0.32 **	-0.49 **	-0.43 **	-0.25 **	-0.26 **	-0.30 **	-0.33 **	-0.34 **
FI	-0.52 *	-1.06 **	-0.46 *	-0.63 **	-0.53 *	-0.76 **	-0.38 *	-0.32
FR	-0.12 *	-0.16 **	-0.10 **	-0.05	-0.05 *	-0.10 **	-0.12 *	-0.12 *
DE	-0.07	0.23 +	0.04	-0.06	0.18	-0.11	-0.06	-0.07
IE	-0.32 *	-0.38 *	-0.34 *	-0.55 **	-0.49 **	-0.43 **	-0.23 *	-0.18
IT	-0.24 **	-0.32 **	-0.26 **	-0.33 **	-0.24 *	-0.25 **	-0.22 *	-0.21 *
JP	-0.28 *	-0.14	-0.23	-0.46 **	-0.34	-0.37 **	-0.20	-0.13
KR	-0.65 *	-0.89 **	-0.63 *	-0.53 *	-1.00 **	-0.80 **	-0.56 *	-0.51 *
NO	-1.25 **	-1.10 **	-1.32 **	-1.00 **	-0.46 *	-1.21 **	-1.31 **	-1.36 **
SE	-0.23 **	-0.46 **	-0.33 **	0.01	0.16 +	-0.29 **	-0.20 **	-0.19 *
UK	-0.12 *	-0.25 **	-0.22 **	-0.23 **	-0.29 **	-0.18 **	-0.11 *	-0.12 *
US	-0.13 *	-0.29 **	-0.24 **	-0.12	-0.27	-0.19 **	-0.12 *	-0.11 *

Note: **/++ indicates significance at 95%-level, */+ indicates significance at 68%-level (based on 1000 bootstrap replications). In contrast to results of Section 3, numbers are based on a 6-variate system estimated via OLS, rather than EGLS.

[11] In contrast to the previous section, results are based on a 6-variate specification estimated via OLS and not EGLS since we do not decompose aggregate hours into intensive and extensive margins. Results between the 6+1-variate and 6-variate specifications are very similar (as they should be), as can be seen e.g. in Appendix Figure 7.

order of 4 and shocks identified at the 40-quarter horizon. The second block of the table contains results for alternative lag choices of 1, 2, 8 or 12. We report results for 1 or 2 lags as they reflect the optimal number of lags selected by either the AIC, SIC or HQQ optimal lag-criteria (in all cases but Sweden, where the AIC indicates 3 lags). Results for the more generous lag lengths of 8 or 12 are reported as they may offer a better approximation of the model's infinite order representation. The third block of Table 6 contains results for horizons of identification of 20, 60 and 80 quarters. Overall, 3 countries stand out in the experiments summarized in Table 6. For the results of different lag orders, we see a sign reversal for Sweden (for a lag order of 12) and Germany (for lag orders of 1, 2 or 12). For the different forecast horizons, the only country that depicts a sign reversal in the impact response is Australia. The positive but insignificant impact response obtained under the baseline specification turns negative for lower horizons over which technology shocks are identified. For an additional sensitivity exercise, we deflate the data with the CPI instead of the GDP deflator. The results, depicted in Appendix Figure 13, are qualitatively similar for all countries. Overall, these results reveal some instabilities in Australia, Germany, and Sweden. For the majority of countries, however, the negative impact response of aggregate hours to positive technology shocks is robust to the VAR's lag order, horizon over which shocks are identified, or the choice of the deflator.

3.3.2 Alternative specification choices

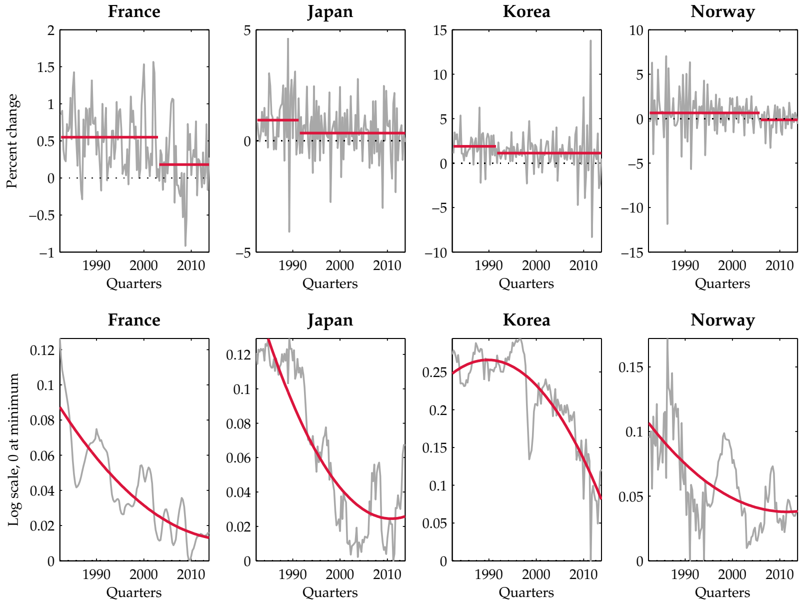
Table 7 reports the size and significance of the median impact response of aggregate hours to positive technology shocks for alternative specifications of labor productivity growth and aggregate hours. Blocks (i) and (ii) contain results for the bivariate specification; blocks (iii) and (iv) are based on our 6-variate specification. In each case, results are reported separately depending on whether labor productivity growth is adjusted for breaks (this is only relevant for the countries highlighted in gray, see Table 15), and depending on how hours enter the system (in first differences (Diff), levels, or quadratically detrended (Det)). The last column of the table shows the results of our baseline specification. Overall, Table 7 shows that the specification choices can matter for the

Table 7: *Impact response of aggregate hours worked to a positive technology shock for different specifications of aggregate hours and labor productivity*

	(i) Bivariate			**(ii) Bivariate, breaks**			**(iii) 6-variate**			**(iv) 6-variate, breaks**		
	Diff	Level	Det	Diff	Level	Det	Diff	Level	Det	Diff	Level	Det
AU	-0.59**	-0.54**	-0.51**	-0.59**	-0.54**	-0.51**	-0.46**	-0.18*	0.01	-0.46**	-0.18*	0.01
AT	-0.34**	-0.35*	-0.31**	-0.34**	-0.35**	-0.31**	-0.42**	-0.33**	-0.21*	-0.42**	-0.33**	-0.21*
CA	-0.29**	-0.41**	-0.47**	-0.29**	-0.41**	-0.47**	-0.33**	-0.29**	-0.32**	-0.33**	-0.29**	-0.32**
FI	-0.98**	-1.03**	-1.29**	-1.05**	-1.15**	-1.52**	-0.60**	-0.49*	-0.36*	-1.25**	-0.35	-0.52*
FR	-0.10**	0.06+	-0.06*	-0.10**	-0.06	-0.06*	-0.08*	0.09	0.05	-0.16**	-0.14	-0.12*
DE	-0.10	0.04	0.01	-0.10	0.04	0.01	-0.15*	-0.05	-0.07	-0.15*	-0.31	-0.07
IE	-0.39**	-0.29	-0.38*	-0.39*	-0.29	-0.38*	-0.20	-0.04	-0.32*	-0.20	-0.04	-0.32*
IT	-0.33**	-0.24*	-0.44**	-0.29**	-0.37**	-0.42**	-0.25**	0.05	-0.04	-0.38**	-0.17*	-0.24**
JP	-0.40**	0.28	-0.32*	-0.45**	-0.24	-0.42**	-0.28**	0.19	-0.17	-0.38**	-0.35**	-0.28*
KR	-1.20**	0.52	-0.98**	-1.43**	-0.25	-0.89**	-1.09**	0.25	-0.32	-1.29**	-0.21	-0.65*
NO	-1.06**	-0.42	-1.13**	-1.24**	-1.22**	-1.35**	0.11	0.35	0.06	-1.16**	-1.24**	-1.25**
SE	-0.42**	-0.50**	-0.43**	-0.42**	-0.49**	-0.43**	-0.44**	-0.29**	-0.23*	-0.44**	-0.29**	-0.23**
UK	-0.34**	-0.08	-0.19*	-0.34**	-0.08	-0.19*	-0.12*	-0.08	-0.12*	-0.12*	-0.08	-0.12*
US	-0.41**	-0.07	-0.06	-0.41**	-0.07	-0.06	-0.13*	-0.11	-0.13*	-0.13*	-0.11	-0.13*

Note: **/++ indicates significance at 95%-level, */+ indicates significance at 68%-level (based on 1000 bootstrap replications). For countries highlighted in gray, trend breaks in labor productivity are found (Table 14).

Figure 2: *Comovement in labor productivity growth and aggregate hours*



Note: Top row reports *quarterly* labor productivity growth; bottom row displays (log) aggregate hours worked, indexed at the minimum of each series.

sign of the impact response. Consider for instance Japan: The median impact response of aggregate hours is positive if hours enter in levels and trend breaks are not accounted for. Removing trend breaks in labor productivity or detrending hours is sufficient to obtain a negative impact response. This finding is similar for France, Korea, and Norway.

Unlike the sensitivity exercises considered in Section 3.3.1, the fact that results are not robust across experiments does not necessarily "raise red flags". Data is detrended and adjusted for breaks prior to estimation specifically because we know—e.g. from Fernald (2007) or Canova et al. (2010)—that estimation results are sensitive to low-frequency comovements in the data. We report results for other specification choices to check whether the issue of low- to medium-frequency movements—shown to be relevant for the US—also matter in other countries. The instabilities in the sign of the median response in aggregate hours apparent for France, Italy, Japan, and Norway are symptomatic of such low-frequency comovements. To visualize this, Figure 2 depicts labor productivity growth (together with our linear trend measure) and the level of aggregate hours (together with our quadratic trend measure). The figure shows that in the four countries considered, the two series have a similar high-low pattern, much like discussed by Fernald (2007) for the US. Since estimated responses switch sign if this comovement is reduced (either by adjusting labor productivity, or aggregate hours, or both), our results reemphasize the importance of controlling for low-frequency patterns. While our results stress the importance of removing long cycles, the choice of using quadratic detrending by contrast to for instance cubic or higher order polynomials is somewhat arbitrary. Our results are robust to these alternative detrending methods.

4 Investment-specific technological change

Up to this point, we have focused on technology shocks that affect the production of all goods homogeneously. We now review and expand these results and consider two distinct sources of technological change, namely disembodied (N) and embodied (I) technology. The work follows Fisher (1999, 2006), Gordon (1990), Greenwood et al. (1988), and Greenwood et al. (2000). We focus on a subsample of 5 countries for which we are able to collect data of sufficient quality.

4.1 Changes to the previous set-up: Data and identification

We start with the amendments to our previous set-up which allow us to disentangle N- and I-shocks. Following the empirical literature on investment-specific technological change, we identify I-shocks from data on the relative price of investment (RPI). An econometric justification for this practice is provided by Schmitt-Grohé and Uribe (2011), who show that the technology transforming consumption into investment goods is approximately linear.[12]

We follow Fisher (2006) and construct data series on the RPI from national accounts statistics as the ratios of chain-weighted deflators for investment and consumption. We define consumption as the sum of consumption in services and in non-durables. Depending on availability, we use either household or private consumption. Investment is the more challenging part to compute: We define it as the sum of consumption in durables, non-residential fixed investment (namely, investment in structures, machinery and equipment, and software) and residential investment (investment in structures and equipment). The exact categories differ across countries depending on availability.[13] We use chain-weighting to aggregate the subcategories into our measures of consumption and investment, as explained in more detail in Appendix Section A.2.3. The construction is based on quarterly data obtained from Statistics Canada, EUROSTAT, the Cabinet Office Japan, and the US Bureau of Economic Analysis, respectively (obtained via Datastream) and covers, as before, the 1982Q3–2013Q4 period.[14] Appendix Section A.2.3 reports selected business cycle moments of the series and depicts the data in (log)-levels and growth rates.

To estimate the VAR including our RPI series, we need to express all series in a common unit. As in Fisher (2006), we use consumption as the common numéraire. To adjust labor productivity, we first express the series in nominal units by multiplying it by its deflator. We then deflate nominal labor productivity by the consumption deflator. Aggregate hours, interest rates, consumption- and investment-to-output ratios, inflation, and the RPI do not need to be adjusted. In particular, the RPI is, by construction, expressed in consumption units, since the price of investment is deflated by

[12] A linear transformation technology is sufficient to ensure that the RPI is exogenous and only depends on investment-specific technology shocks. With non-linear technology, the RPI would instead depend on the amount of resources devoted to investment.

[13] For France, Japan, and the UK, we collect data for gross fixed capital formation (GFCF) by type (namely dwellings, other buildings and structures, transport equipment and other machinery and equipment). For the UK, price increases in residential investment in dwellings are over 400% between 1980 and 2010 and we choose to exclude it. For Canada, we collect the relevant GFCF data for businesses and government. For the US, we use data on private residential and nonresidential investment and government investment.

[14] Our obtained measures of the RPI likely underestimate the rate of investment-specific technological change due to lack of quality adjustment (see e.g. Gordon (1990), Schmitt-Grohé and Uribe (2011) or Cummins and Violante (2002)). For the US, this issue is likely small, as the share of quality-adjusted equipment goods in the national accounts series we use becomes relatively large after 1982. However, there is no reason to assume the situation to be similar in other countries. To our knowledge, information on quality adjustment procedures is not available at an international level, and there are no clear quality adjustment rules across national statistical offices. This drawback in our RPI series is the main reason we report results for the one-technology assumption separately, since the underlying data for Section 3 is likely of much better quality.

the consumption deflator. We use the same baseline specification as before, but now include the RPI in the first position. We now have $Y_t = [\Delta RPI_t, \Delta lp_t, \hat{H}_t, cy_t, iy_t, i_t, \pi_t, \hat{n}_t]'$. Data pre-treatment is as explained in Section 2. In particular, aggregate hours are quadratically detrended and enter in levels. Labor productivity and now also the RPI enter in first differences and are adjusted for trend breaks. Results for applying the endogenous break tests of Bai and Perron (1998, 2003) are summarized in Appendix Table 16.

Regarding the identification, we follow Fisher (2006) and use the two identifying assumptions that (1) I-shocks are the sole source of permanent changes in the RPI and (2) only N- and I-shocks affect labor productivity in the long run. Fisher (2006) shows how these restrictions follow directly from a neoclassical growth model with investment-specific technology. The two identifying assumptions readily translate into the maxFEV set-up. Now, identification follows a 2-step procedure: First, the I-shock is identified as the shock that explains the maximum fraction of forecast error variance of the RPI at a 10-year horizon (following the steps outlined in 2.1). Conditional on having identified the I-shock, the N-shock is identified as the orthogonal shock explaining the maximum fraction of forecast error variance of labor productivity at the 10-year horizon.

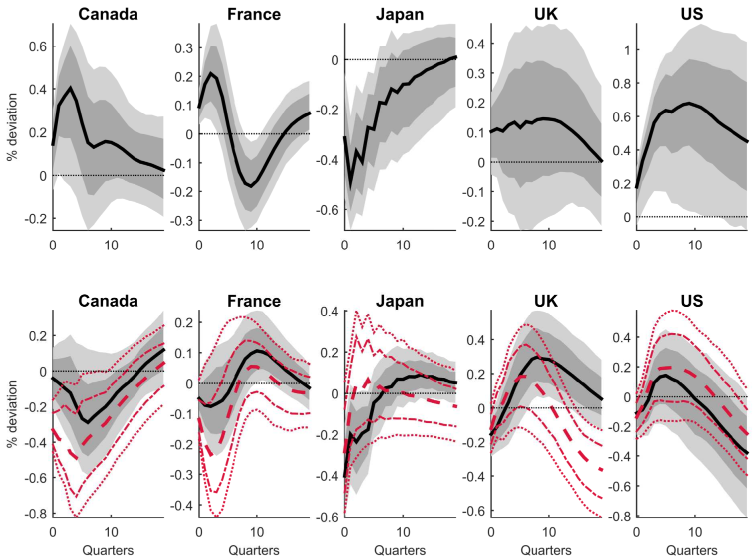
4.2 Results

Figure 3 shows the median response of aggregate hours to one-standard-deviation positive I-shocks (top row) and N-shocks (bottom row) together with 5th, 16th, 84th, and 95th percentiles. Results of the 7+1-specification are depicted in black. For comparison purposes, the bottom row also depicts (in red) the responses of aggregate hours identified under the one-technology assumption for our 6+1-baseline case—these correspond to the results reported in Figure 1. Focusing first on N-shocks, the figure shows similar dynamics under both the one- and two-technology assumptions. For our subsample of five countries, the dynamic responses of aggregate hours appear overall robust to disentangling embodied and disembodied technological change. For our 7+1-specification, the median impact of aggregate hours to positive N-shocks is negative in all cases. At the 68%-level, the impact fall is statistically significant in 4 out of 5 countries. As for the response to I-shocks, Figure 3 depicts a positive median impact response of aggregate hours in 4 countries. The single exception is Japan. For all countries, the estimated impact responses are significant at the 68% level.[15]

Table 9 summarizes the fraction of FEV of aggregate hours and labor productivity explained by N- and I-shocks, respectively. The results are also depicted graphically in Appendix Figure 15. According to the table, I-shocks and N-shocks are roughly equally important for aggregate hours. Further, N-shocks are twice as important as I-shocks in explaining the forecast error variance of labor productivity (at any horizon and for any country). As to the total importance of technology shocks, Table 9 shows that between 15–43% of the FEV of aggregate hours over the first eight years is explained by the two technology shocks. Except for Canada, these numbers are above the ones reported for the one-technology case (Table 2). The same holds for labor productivity: Canada is

[15] We also evaluate the robustness of these findings. By contrast to the dynamics to N-shocks, the response of aggregate hours to I-shocks is not robust across alternative specifications. For instance, Appendix Figure 14 depicts results of a 3+1-specification with $Y_t = [\Delta RPI_t, \Delta lp_t, \Delta H_t, \Delta n_t]$. The figure depicts a fall in aggregate hours in four countries (France is the single exception). Hence, France and Japan are the only countries for which going from a 3+1- to a 7+1-specification does not lead to a sign reversal in the estimated response of aggregate hours to I-shocks. This result is in line with Figure 1, which shows that going from a minimum-size system to our 6+1-baseline specification has large effects on the results of Canada, the UK, and the US.

Figure 3: *Dynamic response of aggregate hours to positive investment-specific shock (top row) and positive neutral technology shock (bottom row)*



Note: Black lines with shaded bands correspond to the 7+1-specification. Dashed red lines depict results of the 6+1-baseline from the previous section. Confidence bands are obtained with 1000 bootstrap replications and cover 68% and 90%. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Table 8: *Fraction of FEV of aggregate hours and labor productivity accounted for by I- and N-shocks*

		Aggregate hours					**Labor productivity**				
		0	**4**	**8**	**16**	**32**	**0**	**4**	**8**	**16**	**32**
N-shock	CA	3	5	8	8	10	26	21	20	20	20
	FR	4	4	7	8	8	32	26	24	24	24
	JP	19	13	13	12	13	51	34	32	32	32
	UK	16	7	8	12	12	66	49	46	46	45
	US	26	11	8	10	14	44	32	31	30	30
I-shock	CA	12	18	16	16	15	3	11	12	13	13
	FR	13	14	13	14	16	15	16	16	16	16
	JP	8	29	30	31	30	17	17	17	17	17
	UK	7	12	13	14	14	3	7	8	8	8
	US	8	9	8	11	14	3	9	10	11	12

Table 9: *Median impact response of employment and hours per employee to positive N- and I-shock*

	N-shocks				**I-shocks**			
Country	n_t		h_t		n_t		h_t	
CA	-0.06	*	0.02		0.07	++	0.10	
FR	0.00		-0.06	**	0.00		0.06	++
JP	-0.04	**	-0.32	**	-0.08	**	-0.14	**
UK	-0.02	**	-0.14	*	0.01		0.07	++
US	-0.08	**	-0.05		0.11	++	0.09	

Note: **/++ indicates significance at 95%-level, */+ indicates significance at 68%-level (based on 1000 bootstrap replications).

Table 10: *Variability in aggregate hours due to fluctuations in employment and hours per employee conditional on N- and I-shocks*

	Mean	CA	FR	JP	UK	US
Conditional on N-shocks						
n_t	55	66	54	38	64	55
h_t	23	9	33	47	10	15
$Cov(h_t, n_t)$	22	25	12	15	27	29
Conditional on I-shocks						
n_t	53	64	31	25	92	52
h_t	28	7	34	70	20	12
$Cov(h_t, n_t)$	19	29	35	5	-12	35

Note: Numbers are based on the variance decomposition (in logs): $\text{Var}(H_t) = \text{Var}(h_t) + \text{Var}(n_t) + 2\text{Cov}(h_t, n_t)$. Cyclical components are extracted via HP-filter with $\lambda = 1600$.

the only country for which the inclusion of investment-specific change does not increase the overall role of technology shocks for labor productivity.

Table 9 summarizes the fraction of FEV of aggregate hours and labor productivity explained by N- and I-shocks, respectively. The results are also depicted graphically in Appendix Figure 15. According to the table, I-shocks and N-shocks are roughly equally important for aggregate hours. Further, N-shocks are twice as important as I-shocks in explaining the FEV of labor productivity (at any horizon and for any country). As to the total importance of technology shocks, Table 9 shows that between 15–43% of the FEV of aggregate hours over the first eight years is accounted for by the two technology shocks. Except for Canada, these numbers are above the ones reported for the one-technology case (Table 2). The same holds for labor productivity: Canada is the only country for which the inclusion of investment-specific change does not increase the overall role of technology shocks for labor productivity.

As to the relative roles of intensive and extensive margins, Table 10 shows the decomposition of the median impact response of aggregate hours to N-shocks and I-shocks in terms of hours per employee and employment. Table 8 decomposes the variability in the conditional components of aggregate hours in terms of hours per employee, employment, or their covariance. Overall, both decompositions do not reveal any clear differences in the roles of the two margins in accounting for the aggregate hours response to either N- or I-shocks. For instance, in Table 10, the relative roles of intensive and extensive margins are largely similar across the two shocks (except Canada). The same holds for Table 8: Whether we condition on N- or I-shocks does not have a strong effect on the relative role of intensive and extensive margins, except France. For France, the relative role of the two labor supply margins switches depending on the source of the shock: While employment accounts for a higher fraction of cyclical fluctuations in aggregate hours conditional on N-shocks, hours per employee are more important conditional on I-shocks. Overall, however, what is most striking from the results are the large cross-country differences in the overall role of intensive and extensive labor supply margins, as discussed in Section 3.2.

5 Concluding remarks

Is the reaction of aggregate hours to technology shocks similar across OECD countries? Based on quarterly data for 14 OECD countries over the 1982Q3-2013Q4 period, we find that it is: Many qualitative features documented for the US hold across the 14 OECD countries of our sample. We observe a negative impact response of aggregate hours to positive permanent technology shocks in 13 out of 14 countries, leading in most instances to a negative comovement between the technology-driven components of labor productivity and aggregate hours. As in Fernald (2007), our estimation results are robust across a wide number of alternative specifications considered, as long as low-frequency comovements between labor productivity and aggregate hours are removed. For a subsample of 5 countries for which we were able to collect data of sufficient quality, the identified dynamics to technology shocks are also shown to be robust to disentangling embodied and disembodied technological change.

While the dynamics in aggregate hours appear similar across countries, our decomposition analysis shows that the intensive and extensive labor supply margins respond differently to technology shocks. In the Anglo-Saxon economies, variation in employment explains the bulk of the aggregate hours response to technology shocks. The opposite holds for Austria, France, Japan, Korea, and Norway. This finding implies that the use of employment data as a proxy for aggregate hours can be misleading. It also emphasizes the importance of labor market rigidities—a topic further explored in the next chapter.

References

Andolfatto, D. (1996). Business cycles and labor-market search. *The American Economic Review*, *86*(1), 112–132.

Bai, J. & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, *66*(1), 47–78.

Bai, J. & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1–22.

Basu, S., Fernald, J. G., & Kimball, M. S. (2006). Are technology improvements contractionary? *The American Economic Review*, *96*(5), 1418–1448.

Benati, L. (2007). Drift and breaks in labor productivity. *Journal of Economic Dynamics and Control*, *31*(8), 2847–2877.

Bernanke, B. & Blinder, A. (1992). The Federal Funds Rate and the channels of monetary transmission. *The American Economic Review*, *82*(4), 901–921.

Blanchard, O. J. & Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review*, *79*(4), 655–673.

Caldara, D., Fuentes-Albero, C., Gilchrist, S., & Zakrajsek, E. (2013). On the identification of financial and uncertainty shocks. 2013 Meeting Papers 965, Society for Economic Dynamics.

Canova, F., Lopez-Salido, D., & Michelacci, C. (2010). The effects of technology shocks on hours and output: a robustness analysis. *Journal of Applied Econometrics*, *25*(5), 755–773.

Chaudourne, J., Fève, P., & Guay, A. (2014). Understanding the effect of technology shocks in SVARs with long-run restrictions. *Journal of Economic Dynamics and Control*, *41*, 154–172.

Christiano, L. J., Eichenbaum, M., & Vigfusson, R. (2003). What happens after a technology shock? NBER Working Papers.

Christiano, L. J., Eichenbaum, M., & Vigfusson, R. (2004). The response of hours to a technology shock: evidence based on direct measures of technology. *Journal of the European Economic Association*, *2*(2-3), 381–395.

Cummins, J. G. & Violante, G. L. (2002). Investment-specific technical change in the united states (1947–2000): measurement and macroeconomic consequences. *Review of Economic Dynamics*, *5*(2), 243–284.

Dupaigne, M. & Fève, P. (2009). Technology shocks around the world. *Review of Economic Dynamics*, *12*(4), 592–607.

Faust, J. & Leeper, E. M. (1997). When do long-run identifying restrictions give reliable results? *Journal of Business & Economic Statistics*, *15*(3), 345–53.

Fernald, J. G. (2007). Trend breaks, long-run restrictions, and contractionary technology improvements. *Journal of Monetary Economics*, *54*(8), 2467–2485.

Fisher, J. D. (1999). The new view of growth and business cycles. *Economic Perspectives*, *23*(1), 35.

Fisher, J. D. (2006). The dynamic effects of neutral and investment-specific technology shocks. *Journal of Political Economy*, *114*(3), 413–451.

Francis, N., Owyang, M. T., Roush, J. E., & DiCecio, R. (2014). A flexible finite-horizon alternative to long-run restrictions with an application to technology shocks. *Review of Economics and Statistics*, *96*(4), 638–647.

Francis, N. & Ramey, V. A. (2005). Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited. *Journal of Monetary Economics*, *52*(8), 1379–1399.

Francis, N. & Ramey, V. A. (2006). The source of historical economic fluctuations: an analysis using long-run restrictions. In *NBER International Seminar on Macroeconomics 2004* (pp. 17–73). NBER Chapters. National Bureau of Economic Research.

Francis, N. & Ramey, V. A. (2009). A century of work and leisure. *American Economic Journal. Macroeconomics*, *1*(2), 189.

Galí, J. (1999). Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations? *The American Economic Review*, *89*(1), 249–271.

Galí, J. (2004). On the role of technology shocks as a source of business cycles: some new evidence. *Journal of the European Economic Association*, *2*(2-3), 372–380.

Galí, J., López-Salido, J. D., & Vallés, J. (2003). Technology shocks and monetary policy: assessing the Fed's performance. *Journal of Monetary Economics*, *50*(4), 723–743.

Glocker, C. & Wegmueller, P. (2018). International evidence of time-variation in trend labor productivity growth. *Economics Letters*, *167*, 115–119.

Gordon, R. J. (1990). *The Measurement of Durable Goods Prices*. University of Chicago Press.

Greenwood, J., Hercowitz, Z., & Huffman, G. W. (1988). Investment, capacity utilization, and the real business cycle. *The American Economic Review*, *78*(3), 402–417.

Greenwood, J., Hercowitz, Z., & Krusell, P. (2000). The role of investment-specific technological change in the business cycle. *European Economic Review*, *44*(1), 91–115.

Hansen, G. D. (1985). Indivisible labor and the business cycle. *Journal of Monetary Economics*, *16*(3), 309–327.

Ireland, P. N. (2000). Interest rates, inflation, and Federal Reserve policy since 1980. *Journal of Money, Credit and Banking*, *32*(3), 417–34.

Justiniano, A., Primiceri, G., & Tambalotti, A. (2011). Investment shocks and the relative price of investment. *Review of Economic Dynamics*, *14*(1), 101–121.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions. *Review of Economics and Statistics*, *80*(2), 218–230.

Kydland, F. E. & Prescott, E. (1982). Time to build and aggregate fluctuations. *Econometrica*, *50*(6), 1345–1370.

Long, J., John B & Plosser, C. I. (1983). Real business cycles. *Journal of Political Economy*, *91*(1), 39–69.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer Science & Business Media.

Merz, M. (1995). Search in the labor market and the real business cycle. *Journal of Monetary Economics*, *36*(2), 269–300.

Mortensen, D. T. & Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The Review of Economic Studies*, *61*(3), 397–415.

Ohanian, L. E. & Raffo, A. (2012). Aggregate hours worked in OECD countries: new measurement and implications for business cycles. *Journal of Monetary Economics*, *59*(1), 40–56.

Schmitt-Grohé, S. & Uribe, M. (2011). Business cycles with a common trend in neutral and investment-specific productivity. *Review of Economic Dynamics*, *14*(1), 122–135.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: the effects of monetary policy. *European Economic Review*, *36*(5), 975–1000.

Smets, F. & Wouters, R. (2007). Shocks and frictions in US business cycles: A bayesian DSGE approach. *The American Economic Review*, *97*(3), 586–606.

Taylor, J. B. (1993). Discretion versus policy rules in practice. In *Carnegie-Rochester Conference Series on Public Policy* (Vol. 39, pp. 195–214). Elsevier.

Uhlig, H. (2003). What drives GNP? Unpublished manuscript, Euro Area Business Cycle Network.

Uhlig, H. (2004). Do technology shocks lead to a fall in total hours worked? *Journal of the European Economic Association*, *2*(2-3), 361–371.

A Appendix

A.1 Empirical methodology

A.1.1 Estimated generalized least squares (EGLS)

For the decomposition of the aggregate hours response into intensive and extensive margins of labor supply, we append information on employment in the last position of our VAR and use generalized least squares to estimate a SVAR with linear constraints, following Lütkepohl (2005). To introduce the methodology, it is convenient to rewrite the system in (2.1) in compact form,

$$X = BZ + U, \tag{A.1}$$

where $X := [Y_1, \ldots, Y_T]$, $Z := [Z_0, \ldots, Z_{T-1}]$ for $Z_t := [1, Y_t, \ldots, Y_{t-4+1}]'$, $B := [B_0, \ldots, B_4]$, and $U := [u_1, \ldots, u_T]$. We can then impose any linear restriction of interest on B by writing the least-squares (LS) estimator as $\boldsymbol{\beta} := vec(B) = R\boldsymbol{\gamma} + r$. With p denoting the lag-order of the system (in our standard case $p = 4$), R is a known $((n^2p + n) \times M)$ matrix of rank M, $\boldsymbol{\gamma}$ is a $M \times 1$ vector of unrestricted unknown parameters, and r is a vector of normalizing constants with dimension $((n^2p + n) \times 1)$. In our specific case, r is a vector of zeros and R is specified such that the desired constraints on B_j hold, with M amounting to $(n^2p - (n-1)p)$. With R known, the VAR model can be expressed as

$$\boldsymbol{Y} := vec(X) = \left(Z' \otimes I_n\right) \boldsymbol{\beta} + \boldsymbol{u} = \left(Z' \otimes I_n\right) \left(R\boldsymbol{\gamma} + r\right) + \boldsymbol{u}, \tag{A.2}$$

where $\boldsymbol{u} := vec(U)$. Defining $\boldsymbol{z} := \boldsymbol{Y} - (Z' \otimes I_n)\, r$, expression (A.2) simplifies to

$$\boldsymbol{z} = \left(Z' \otimes I_n\right) R\boldsymbol{\gamma} + \boldsymbol{u}. \tag{A.3}$$

Lütkepohl (2005) shows that a generalized LS (GLS) estimator can be calculated as

$$\hat{\boldsymbol{\gamma}} = \left[R'(ZZ' \otimes \Sigma_u^{-1})R\right]^{-1} R'(Z \otimes \Sigma_u^{-1})\boldsymbol{z}. \tag{A.4}$$

The matrix Σ_u is typically unknown and has to be estimated from the data. We use the consistent and unbiased estimator from the unrestricted VAR, $\hat{\Sigma}_u = \frac{1}{T-np-1}\hat{U}\hat{U}'$, to obtain the EGLS estimator, $\hat{\hat{\boldsymbol{\gamma}}}$, which has the same asymptotic properties as the GLS estimator $\hat{\boldsymbol{\gamma}}$:

$$\hat{\hat{\boldsymbol{\gamma}}} = \left[R'(ZZ' \otimes \bar{\Sigma}_u^{-1})R\right]^{-1} R'(Z \otimes \bar{\Sigma}_u^{-1})\boldsymbol{z}. \tag{A.5}$$

Finally, the estimator for $\boldsymbol{\beta}$ is given by $\hat{\hat{\boldsymbol{\beta}}} = R\hat{\hat{\boldsymbol{\gamma}}} + r$.

A.1.2 Identification: Maximum fraction of forecast error variance approach

The following explanations are drawn from Francis et al. (2014), Caldara, Fuentes-Albero, Gilchrist, and Zakrajsek (2013) and Uhlig (2003). To provide some structure on the methodology, it is helpful to rewrite equation (2.1) into the moving-average representation or its equivalent representation in terms of the lag polynomial $\Phi(L)$,

$$Y_t = \sum_{s=0}^{\infty} \Phi_s A \epsilon_{t-s} \equiv \Phi(L) A \epsilon_t, \tag{A.6}$$

where $\Phi(L) \equiv I + \Phi_1 L + \Phi_2 L^2 + \ldots$. The version of the maxFEV approach we consider is to identify a shock by searching for innovations that explain the maximum amount of FEV of a specified variable at some target horizon k. We are hence interested in the k-step ahead forecast error of equation (A.6), which is given by:

$$Y_{t+k} - \mathbb{E}_t[Y_{t+k}] = \sum_{s=0}^{k-1} \Phi_s A \epsilon_{t+k-s}. \tag{A.7}$$

With labor productivity being ordered first in the SVAR, the k-step ahead forecast error due to the technology shock is given by

$$y_{1,t+k} - \mathbb{E}_t[y_{1,t+k}] = e_1 \; [\sum_{s=0}^{k-1} \Phi_s A \epsilon_{t+k-s}], \tag{A.8}$$

where e_1 has dimension $1 \times n$ and takes the form $e_1 = [1 \quad 0 \quad ... \quad 0]$. Following Caldara et al. (2013), what is needed to find the maximum fraction of FEV is an orthogonal matrix Q such that $A = \widetilde{A}Q$, where $\widetilde{A}$ denotes the Cholesky decomposition of Ω. Then, finding the innovation that accounts for the maximum fraction of FEV of the first variable in Y_t amounts to finding the first column of Q (denoted by q_1) by solving:

$$\max_{q_1} \quad e_1 \; [\sum_{s=0}^{k-1} \Phi_s \widetilde{A} q_1 q_1' \widetilde{A}' \Phi_s'] \; e_1' = q_1' S q_1 \tag{A.9}$$

$$\text{subject to} \quad q_1' q_1 = 1.$$

The solution to equation (A.9) corresponds to finding q_1—the eigenvector of S of the largest eigenvalue λ.

A.1.3 Identification: Long-run restrictions

Equations (2.1) and (2.2) imply that the long-run impact of structural shocks (LRI) is given by

$$LRI = \underbrace{[I_2 - B(1)]^{-1} A_0}_{CA_0} \tag{A.10}$$

in which $B(1) = B_1 L - B_2 L^2 - B_3 L^3 - B_4 L^4$. With ΔLP_t being the first variable in the vector Y_t and x denoting non-zero elements, the matrix of long-run impacts has to take the form:

$$[I_2 - B(1)]^{-1} A_0 = \begin{bmatrix} x & 0 \\ x & x \end{bmatrix}. \tag{A.11}$$

Equation (A.11) illustrates that the matrix of long-run impacts is lower triangular, so that the technology shock can be identified by a Cholesky decomposition of the long-run impact matrix.

A.2 Data

A.2.1 Aggregate hours and its components

Table 11 summarizes basic descriptive statistics for aggregate hours per capita and its components for the 1982Q3–2013Q4 period. In the table, data on aggregate hours per capita H_t and per worker h_t is expressed in weekly units; employment is expressed as a percentage of the population aged 16 to 64. The table shows substantial cross-country heterogeneity in the labor supply indicators. Consider for instance France and Korea, which display the lowest and highest weekly median hours per capita. With 19.7 and 31.1 hours respectively, the difference in the median amounts to more than 10 hours per week. Because the employment rate is lower in Korea, the difference is even larger in per worker terms: the median work week in Korea is almost 18 hours larger than in France.

Table 11: *Data summary statistics for aggregate hours worked and its components (1982Q3–2013Q4)*

	Weekly hours worked per capita				**Weekly hours worked per worker**				**Employment per capita (in %)**			
Country	Med	Min	Max	StD	Med	Min	Max	StD	Med	Min	Max	StD
AU	23.4	21.2	24.5	0.73	34.1	32.3	35.0	0.72	68.5	62.1	73.9	3.21
AT	20.1	19.0	22.7	1.06	34.5	31.2	37.9	1.95	58.9	57.8	61.6	0.98
CA	24.2	21.7	25.3	0.83	34.1	32.5	34.9	0.62	71.4	63.6	75.6	3.12
FI	22.7	19.9	26.0	1.58	33.8	31.8	35.8	1.01	69.1	58.5	73.5	4.12
FR	19.7	19.0	21.3	0.50	30.2	28.2	32.5	1.47	65.5	63.0	68.7	2.03
DE	20.3	19.4	21.9	0.69	28.9	26.5	33.2	1.91	70.1	64.7	77.8	3.46
IE	22.3	20.8	25.5	1.37	37.8	34.4	42.9	3.01	60.3	51.3	70.8	6.63
IT	21.6	19.8	22.6	0.65	35.6	33.4	36.4	0.79	60.5	57.7	64.4	1.77
JP	26.8	25.3	28.8	1.15	35.4	32.9	40.7	2.60	74.3	70.1	80.4	2.50
KR	31.1	24.6	33.0	1.82	47.9	36.0	55.4	4.87	65.2	56.5	70.3	3.68
NO	21.5	20.4	24.2	0.73	28.2	26.7	30.9	1.00	77.3	72.5	79.9	2.01
SE	23.2	21.9	24.6	0.63	30.8	29.3	32.2	0.80	75.5	70.3	82.1	3.17
UK	23.4	22.5	24.9	0.57	32.8	31.3	34.6	1.00	71.8	65.4	73.9	2.32
US	25.8	23.5	27.9	1.16	33.1	32.2	33.7	0.32	78.3	71.3	83.3	3.09

Table 12 reports results of Phillips-Perron unit root tests for aggregate hours per capita. When we test the null hypothesis of a unit root in aggregate hours against the alternative of a time-trend, we reject the null for 13 out of 14 countries. Hours per capita appear to be stationary only in France (and in the United Kingdom if one follows the ρ-statistic). When we test the null hypothesis of a unit root in aggregate hours against the alternative of no time-trend, we find a unit root in all series except France.

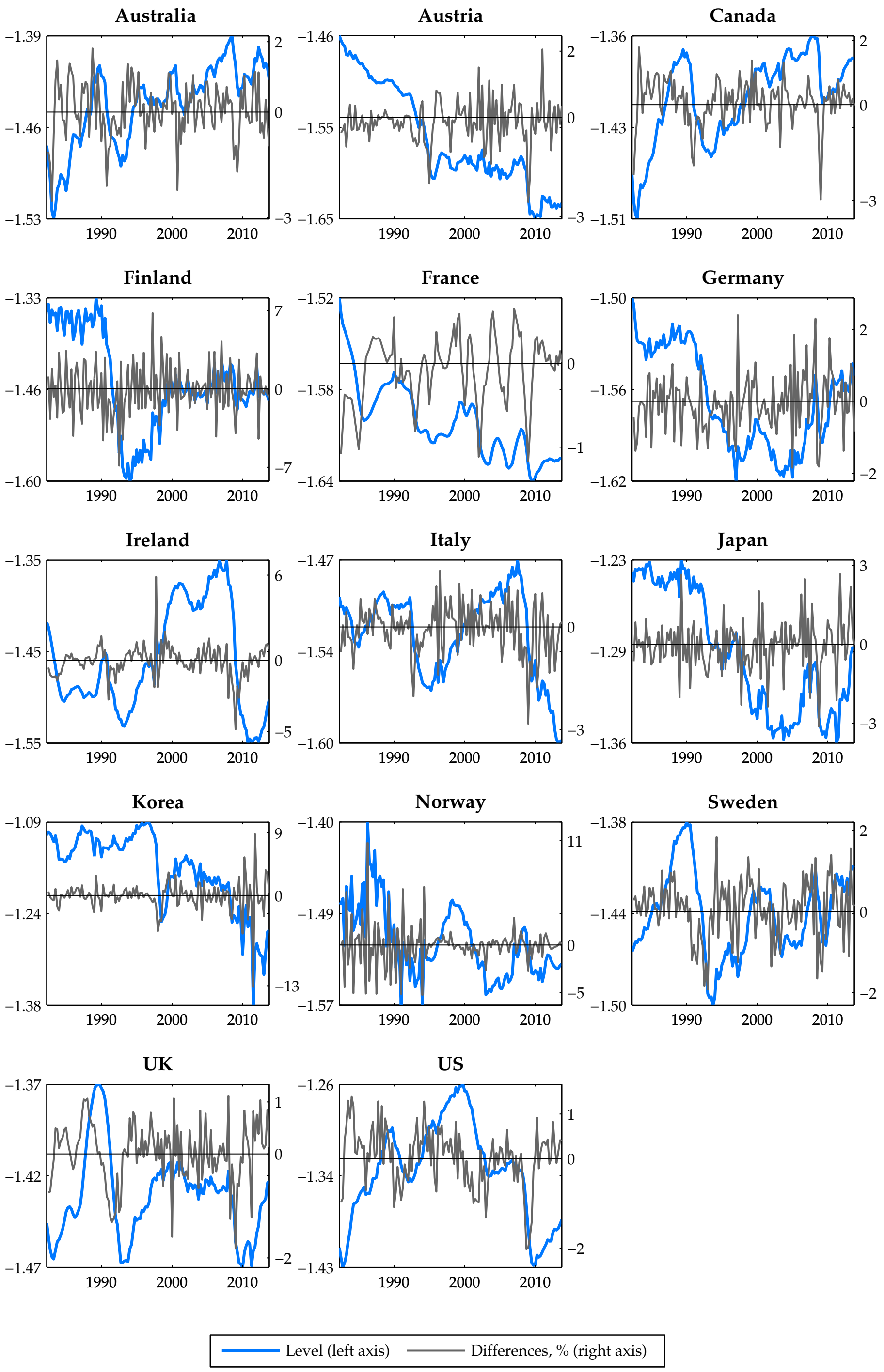
Table 12: *Phillips-Perron-test for unit root*

	Hours without time trend$^{(ii)}$		**Hours with time trend**$^{(iii)}$	
Country$^{(i)}$	ρ-Test$^{(iv)}$	$t_{\hat{\phi}}$-Test$^{(v)}$	ρ-Test$^{(iv)}$	$t_{\hat{\phi}}$-Test$^{(v)}$
AU	-5.43	-1.71	-13.04 *	-2.52
AT	-2.30	-1.61	-9.61	-2.24
CA	-4.62	-1.78	-6.30	-1.84
FI	-5.28	-1.85	-5.57	-1.74
FR	-7.57 **	-3.51 **	-13.59 ***	-3.49 **
DE	-5.23	-2.13	-3.77	-1.24
IE	-2.72	-1.16	-2.83	-1.18
IT	-0.69	-0.25	-1.62	-0.58
JP	-3.40	-1.54	-5.08	-1.14
KR	-6.09	-1.65	-19.35	-3.26
NO	-18.29	-3.33	-31.09	-4.42
SE	-5.81	-1.66	-5.82	-1.66
UK	-5.23 **	-1.64	-5.40	-1.68
US	-2.71	-1.36	-3.27	-1.72

Note: (i) ISO country codes; (ii) Neither theory nor visual inspection of plots in Figure 4 provides clear guidance of whether the econometrician should include a time trend in the unit root test; (iii) As hours per capita are bounded above and below, the inclusion of a time trend is purely dependent on the considered time window; (iv) Statistic is $T(\hat{\rho}-1)$ with T being the sample length and $\hat{\rho}$ the OLS estimate of $y_t = \rho y_{t-1} + u_t$; (v) OLS t-test statistic for null hypothesis that $\rho = 1$. ***, ** and * imply that we fail to reject the null of a unit root at 1%, 5% and 10% significance, respectively. p-values are computed based on 1'000 bootstrap replications.

Figure 4 plots the series for total hours per capita in both log-levels and first differences and shows that the heterogeneity across countries is not just in the levels. The different developments of total hours across countries hint at substantial diversity in the national demographic and social (e.g. educational) trends. For instance, the total hours series of Austria or France depict an almost linear fall in per-capita hours worked compared to an inverse U-shaped pattern in the US.

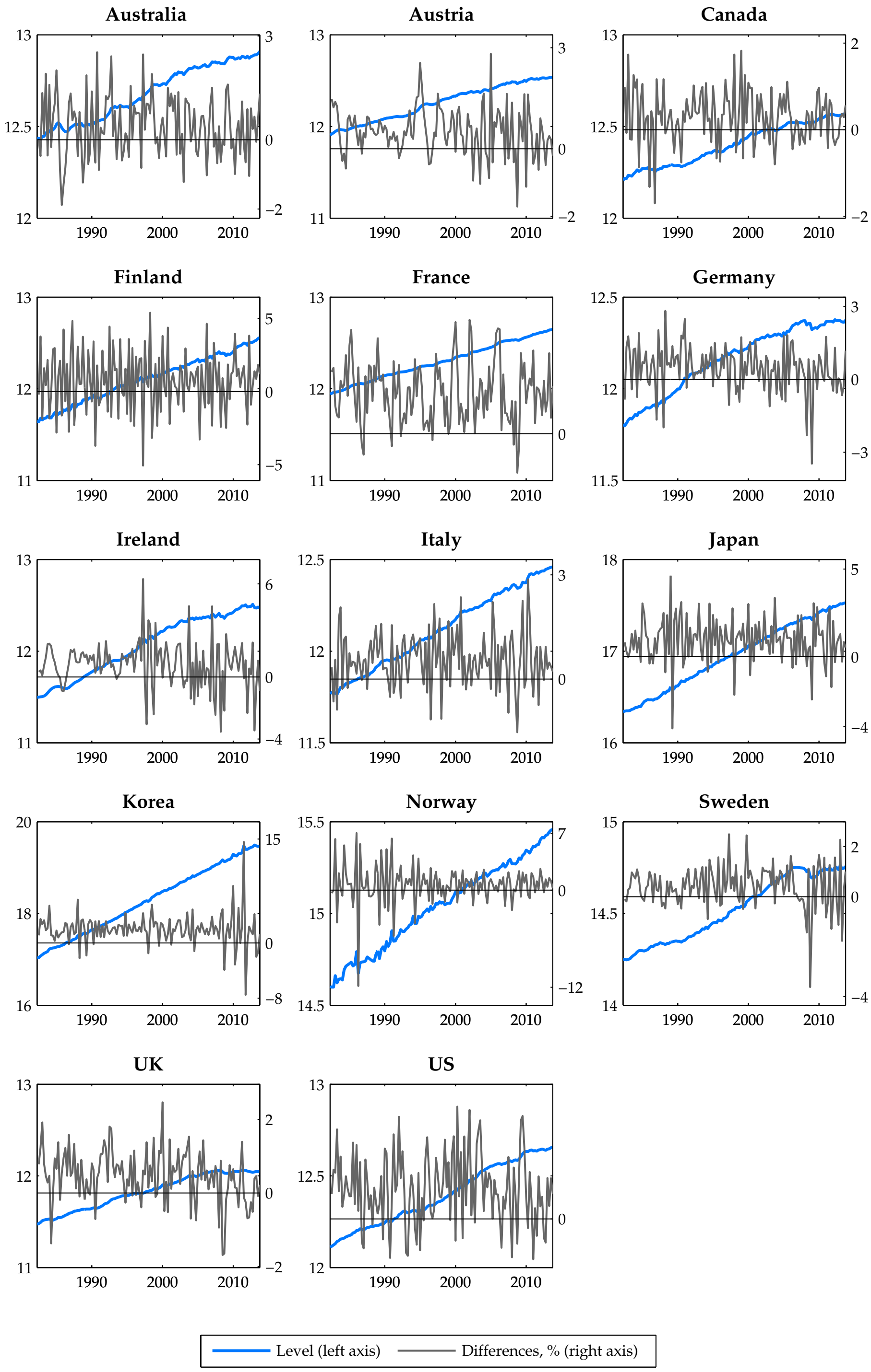
Figure 4: *Aggregate per capita hours in log-levels and first differences*



A.2.2 Labor productivity

Figure 5 displays the log-level of labor productivity together with the quarter-by-quarter growth rates. Both the volatility and mean rate of growth display substantial heterogeneity across countries and over time. For instance, for Germany and Italy, the figure depicts a substantial reduction in the average growth rates of labor productivity in the mid-1990s.

Figure 5: *Labor productivity in log-levels and first differences*



A.2.3 The relative price of investment

Computation: In the following we explain the the chain-weighting methodology applied to compute the consumption deflator. An investment deflator is computed in an analog way. The ratio between the investment and consumption deflators then gives a measure of the real price of investment.

Computing the chain-weighted growth in the consumption deflator involves making two calculations of growth for each period. The only difference is the base period used for the quantities: once, we compute the growth in the prices by using the quantities of the period itself, and once by using the quantities of the preceding period.[16] With Q denoting quantities and P the deflator, we compute:

$$gr_1 = \frac{\Sigma_i P_{t,i} Q_{t,i}}{\Sigma_i P_{t-1,i} Q_{t,i}} = \frac{\Sigma_i C_{curr,t,i}}{\Sigma_i \frac{C_{curr,t-1,i}}{C_{con,t-1,i}} C_{con,t,i}} \tag{A.12}$$

$$gr_2 = \frac{\Sigma_i P_{t,i} Q_{t-1,i}}{\Sigma_i P_{t-1,i} Q_{t-1,i}} = \frac{\Sigma_i \frac{C_{curr,t,i}}{C_{con,t,i}} C_{con,t-1,i}}{\Sigma_i C_{curr,t-1,i}} \tag{A.13}$$

where the indices i denote the different components of consumption used. The right-hand equations show the computations in terms of indices measured in constant prices (C_{con}, giving Q) and current prices (C_{curr}, corresponding to PQ). The chain-weighted growth in the consumption deflator is then computed as the geometric average of the two growth rates, namely:

$$\Delta P_{CW} = (gr_1 \times gr_2)^{0.5} - 1 \tag{A.14}$$

Ideally, we would want chain-weighted indices instead of constant price series. However, such series are in most cases only available after the 1990s. The choice of a fixed year means that one is using a price structure that becomes more and more remote from the current structure the further we move away from the base year. To alleviate somewhat from this problematic, we take several series for different base years and link them via growth rates into one index.

Business cycle moments: Table 13 reports selected business cycle moments for our RPI series. According to the table, the unconditional correlation between the RPI and GDP (measured in consumption units, as explained below) is positive in all countries. The correlation is lowest in the US and highest in Japan. Apart from the UK, we find that the volatility in the RPI is lower than in GDP.

Table 13: *Business cycle moment for the RPI (1982Q3–2013Q4)*

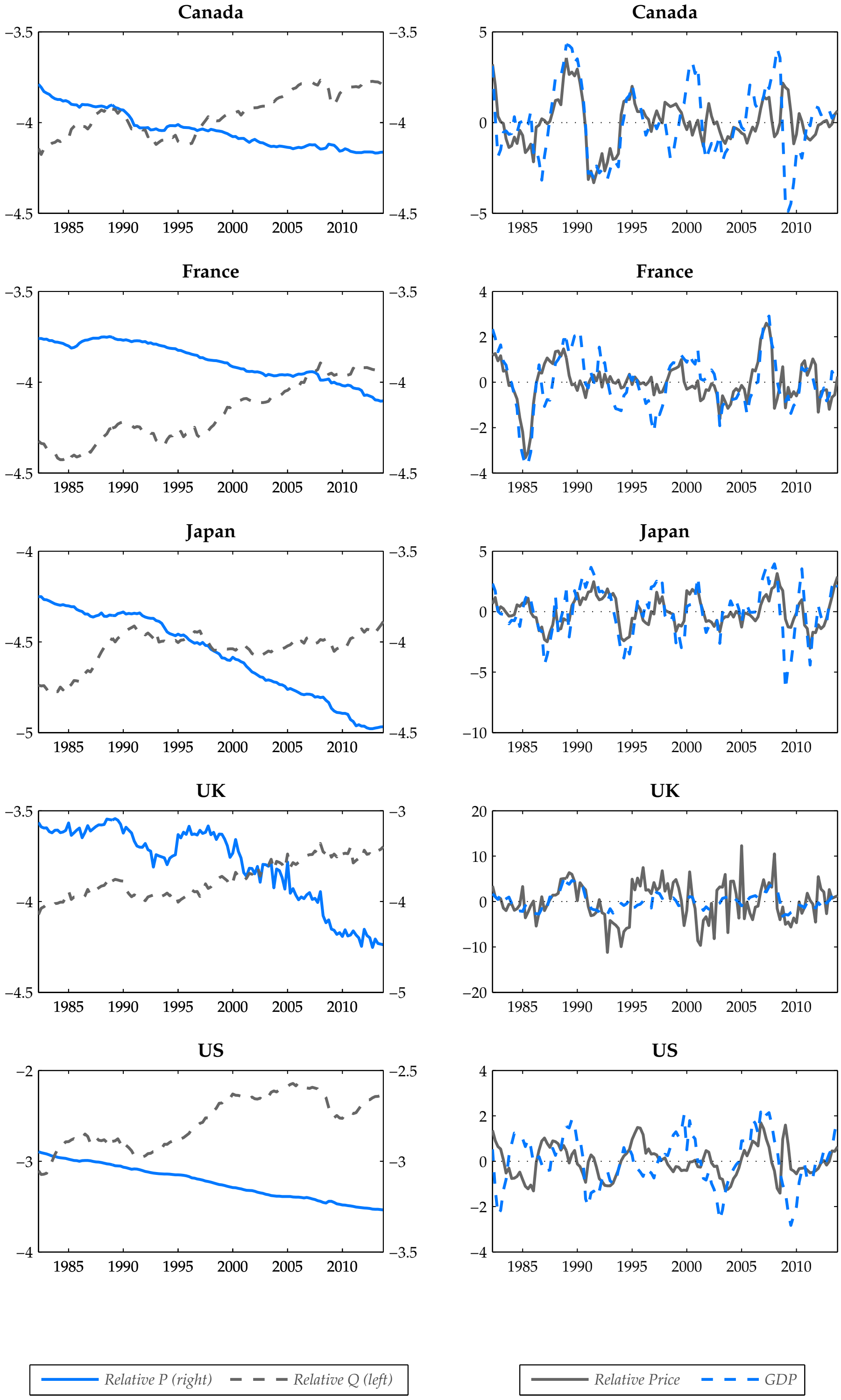
	CA	FR	JP	UK	US
$\rho(RPI, GDP)$	0.504	0.728	0.778	0.533	0.191
σ_{RPI}	1.310	0.917	1.246	4.178	0.706
$\sigma_{RPI}/\sigma_{GDP}$	0.663	0.746	0.624	2.250	0.622

Note: Cyclical components are obtained based on the HP-filter with $\lambda = 1600$. All correlations $\rho(\cdot)$ are significant at the 5% level.

Visual inspection: Figure 6 shows a clear negative trend in the RPI for all five countries. While the decline in the RPI is near monotone in Canada, France, Japan, and the US, there is substantial variation in the UK series. The figure further shows that the large price declines coincide with large increases in the quantity of investment. Investment and its relative price are clearly negatively correlated.

[16] To give an example, to compute the chain-weighted growth rates in 2000 prices, we compute two distinct growth rates with either 1999 or 2000 as base year and take their geometric average.

Figure 6: *Relative price and quantity of investment*



Note: The left column shows the relative price of investment RPI and relative quantity of investment in terms of consumption goods. The right column shows the cyclical components of the RPI, computed by applying the HP-filter with $\lambda = 1600$.

A.2.4 Employment protection indicators

Table 14 depicts the OECD Employment Protection Legislation rankings. A higher value of the indicator corresponds to higher restrictions on firing.

Table 14: *Selected OECD indicators on Employment Protection Legislation for 2013*

	Protection of permanent workers against individual and collective dismissals		**Regulation on temporary forms of employment**	
Country$^{(i)}$	Value$^{(ii)}$	Rank$^{(iii)}$	Value$^{(ii)}$	Rank$^{(iii)}$
AU	1.94	29	1.04	30
AT	2.44	12	2.17	16
CA	1.51	32	0.21	34
FI	2.17	23	1.88	20
FR	2.82	5	3.75	3
DE	2.84	4	1.75	22
IE	2.07	27	1.21	27
IT	2.89	3	2.71	8
JP	2.09	25	1.25	26
KR	2.17	22	2.54	9
NO	2.31	19	3.42	4
SE	2.52	10	1.17	29
UK	1.59	31	0.54	32
US	1.17	33	0.33	33

Note: (i) ISO country codes. (ii) The scale of these values goes from 0 (least restrictions) to 6 (most restrictions). Among all 34 countries, the average value of the two measures reported is 2.27 and 2.07, respectively. (iii) Rank among all 34 member countries.

A.2.5 Break tests

Table 15 summarizes our results of testing for breaks in labor productivity. We follow the same estimation strategy as in Benati (2007). The null hypothesis of no structural break is rejected based on either the *UDmax* and *WDmax* test statistic for Finland, France, Italy, Japan, Korea, and Norway. For all other countries, the double maximum tests indicate that the null of no break cannot be rejected at the 10% significance level. Turning to the number of breaks, results from the $sup - F(l+1|l)$ tests (Appendix Table 15) suggest that no country experienced two structural breaks over our sample period. Further information on the methodology, as well as detailed cross-country evidence on breaks and drifts in labor productivity, are also contained in Glocker and Wegmueller (2018).

Table 15: *Tests for multiple breaks in labor productivity growth*

	Double maximum tests				$sup - F(l+1\|l)$	**Break dates**
Country	UDmax		WDmax		$F(2\|1)$	
AU	19.12		22.17			—
AT	26.29		31.17			—
CA	41.39		47.78			—
FI	21.81	**	24.87	**	17.73	2008Q1 [2003Q4; 2009Q2]
FR	60.93	*	70.73	*	46.06	2003Q1 [2001Q4; 2004Q4]
DE	59.69		69.27			—
IE	49.06		58.37			—
IT	33.90	*	40.00	*	27.15	1995Q4 [1993Q1; 1998Q3]
JP	29.28	**	34.94	**	24.28	1991Q2 [1990Q3; 1993Q3]
KR	17.57	***	19.99	***	15.09	1991Q3 [1985Q2; 1992Q4]
NO	28.13	***	31.54	***	22.05	2005Q4 [2005Q2; 2008Q2]
SE	33.68		39.61			—
UK	53.85	*	62.78			—
US	45.18		51.04			—

Note: ***, ** and * significant at 1%, 5% and 10%, respectively. Break dates show 90% confidence intervals.

Table 16 summarizes our results of testing for breaks in labor productivity expressed in consumption units and in the RPI. For labor productivity, the null hypothesis of no structural break is rejected for Japan and the UK. For the RPI, the null is rejected for Japan.

Note that since labor productivity is now expressed in consumption units, the identified break dates are not the same as reported in Table 15. Given this finding, we have also checked that our results are robust to how labor productivity enters the system. While not shown, our results are robust, *i.e.* they do not crucially depend on the relative dynamics of consumption versus output prices.

Table 16: *Tests for multiple breaks at unknown points in LP and RPI growth*

	Double maximum tests				**Break dates**
Country	UDmax		Wdmax		
LP in consumption units					
CA	29,39		34,51		—
FR	65,60	*	76,33		—
JP	49,59	**	57,90	**	1991Q4 [1990Q4; 1993Q1]
UK	52,43	**	61,26	**	2007Q4 [2006Q1; 2008Q2]
US	41,01		49,47		—
RPI					
CA	54,83		65,33		—
FR	79,63		92,36		—
JP	38,62	**	45,36	**	1991Q4 [1986Q1; 1993Q4]
UK	31,11		35,82		—
US	36,70		41,68		—

Note: ***, ** and * indicate 1, 5, and 10% significance levels. Break dates show 90% confidence intervals.

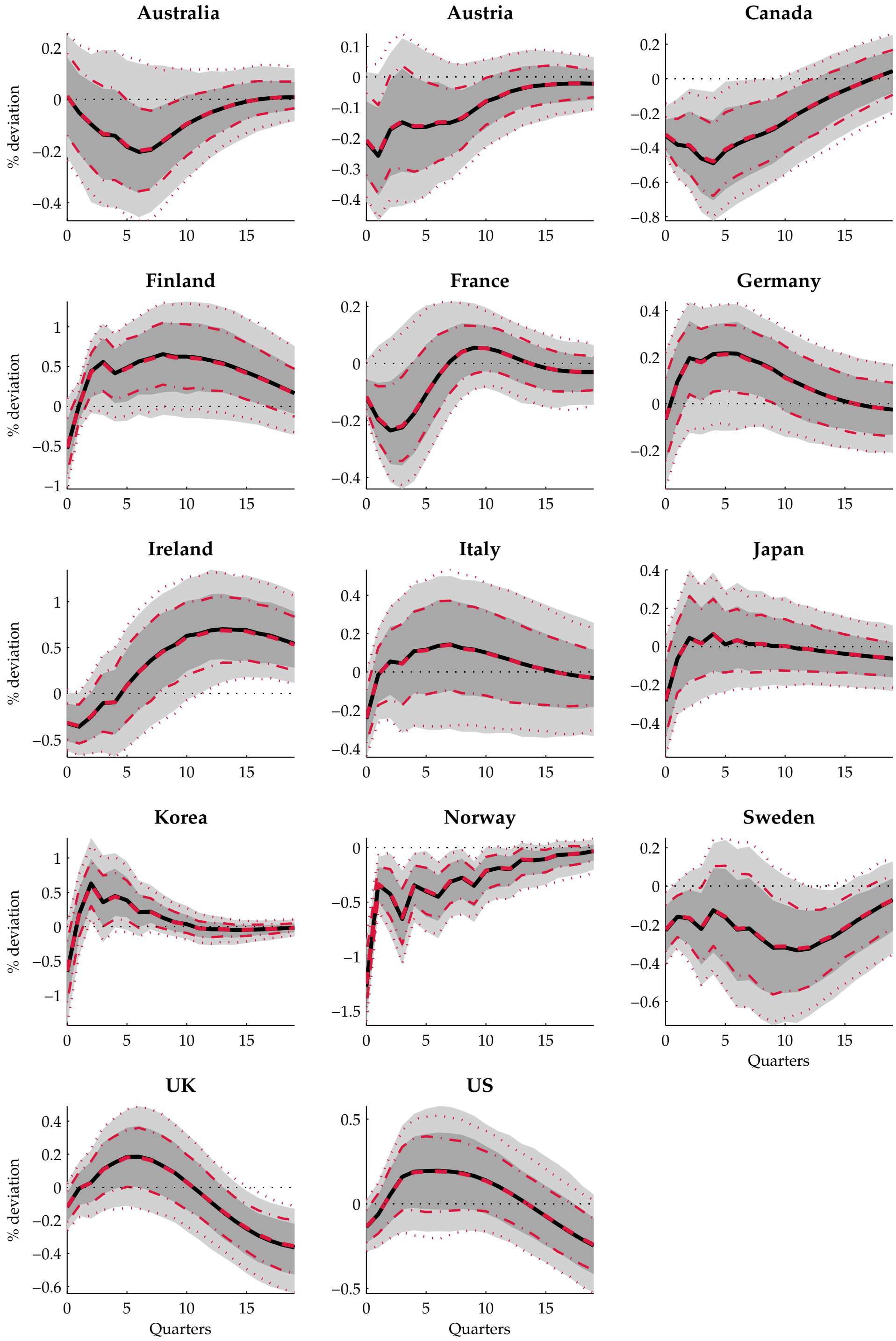
A.3 Additional results

Table 17: *Robustness: Comparison of baseline SVAR to simulations with lag length selection based on different criteria*

	Baseline	AIC		SIC		HQQ	
Country		Response	Lags	Response	Lags	Response	Lags
AU	0.01	0.02	2	-0.05	1	-0.05	1
AT	-0.21 *	-0.27 **	2	-0.33 **	1	-0.33 **	1
CA	-0.32 **	-0.43 **	2	-0.49 **	1	-0.49 **	1
FI	-0.52 *	-0.46 *	2	-1.07 **	1	-1.07 **	1
FR	-0.12 *	-0.10 **	2	-0.10 **	2	-0.10 **	2
DE	-0.07	0.25 +	1	0.25 +	1	0.25 +	1
IE	-0.32 *	-0.36 *	1	-0.36 *	1	-0.36 *	1
IT	-0.24 **	-0.26 **	2	-0.32 **	1	-0.32 **	1
JP	-0.28 *	-0.14	1	-0.14	1	-0.14	1
KR	-0.65 *	-0.89 **	1	-0.89 **	1	-0.89 **	1
NO	-1.25 **	-1.10 **	1	-1.10 **	1	-1.10 **	1
SE	-0.23 **	-0.17 *	3	-0.43 **	1	-0.43 **	1
UK	-0.12 *	-0.25 **	1	-0.25 **	1	-0.25 **	1
US	-0.13 *	-0.24 **	2	-0.29 **	1	-0.24 **	2

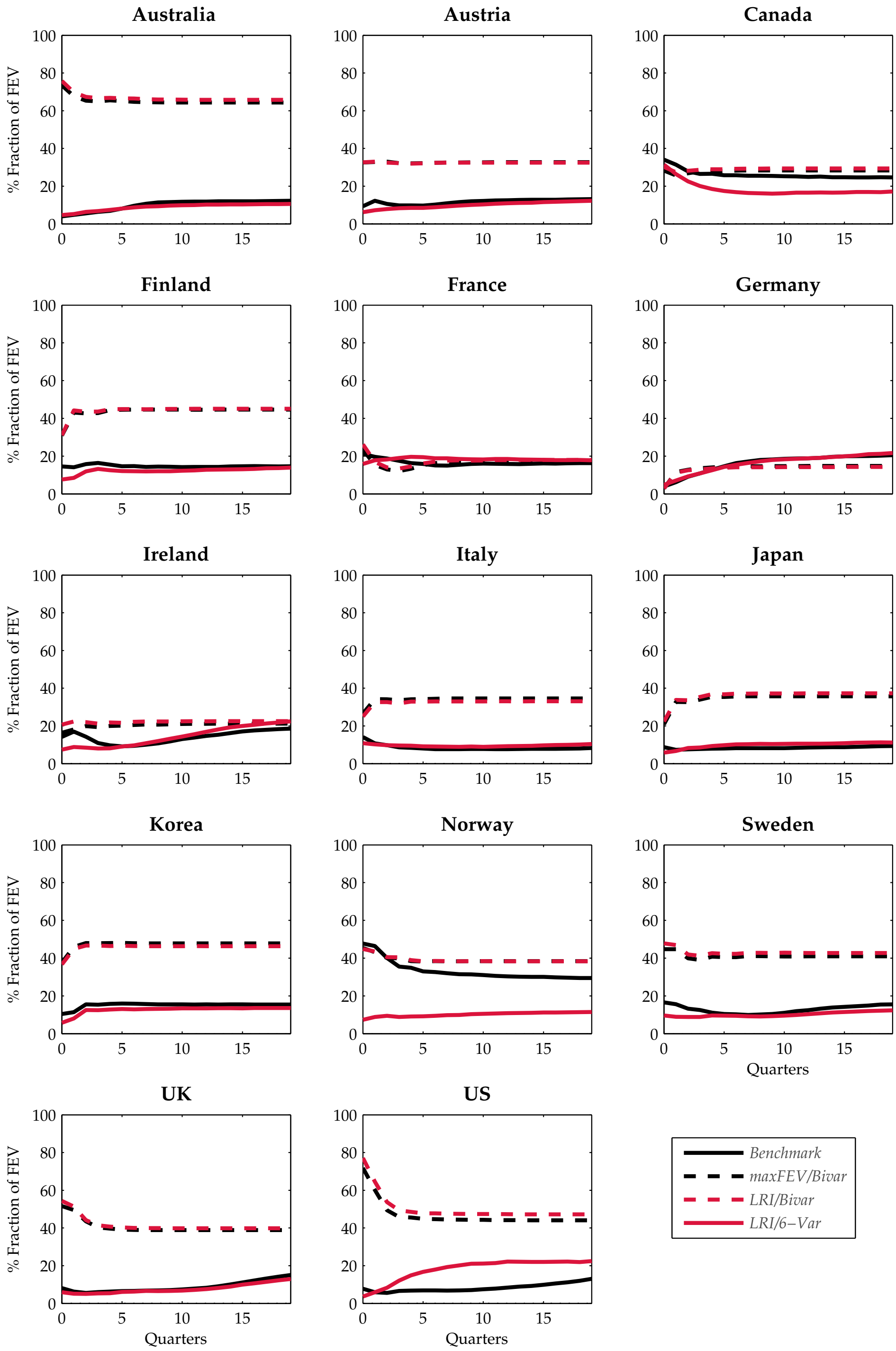
Note: AIC: Akaike Information Criterion; SIC: Schwartz Information Criterion; HQC: Hannan-Quinn Information Criterion. Contemporaneous response of hours worked: ** significant at 95%; * significant at 68%. Results are based on 1000 bootstrap replications. Baseline features 4 lags.

Figure 7: *Dynamic response of aggregate hours to a positive technology shock: Comparing OLS and EGLS*



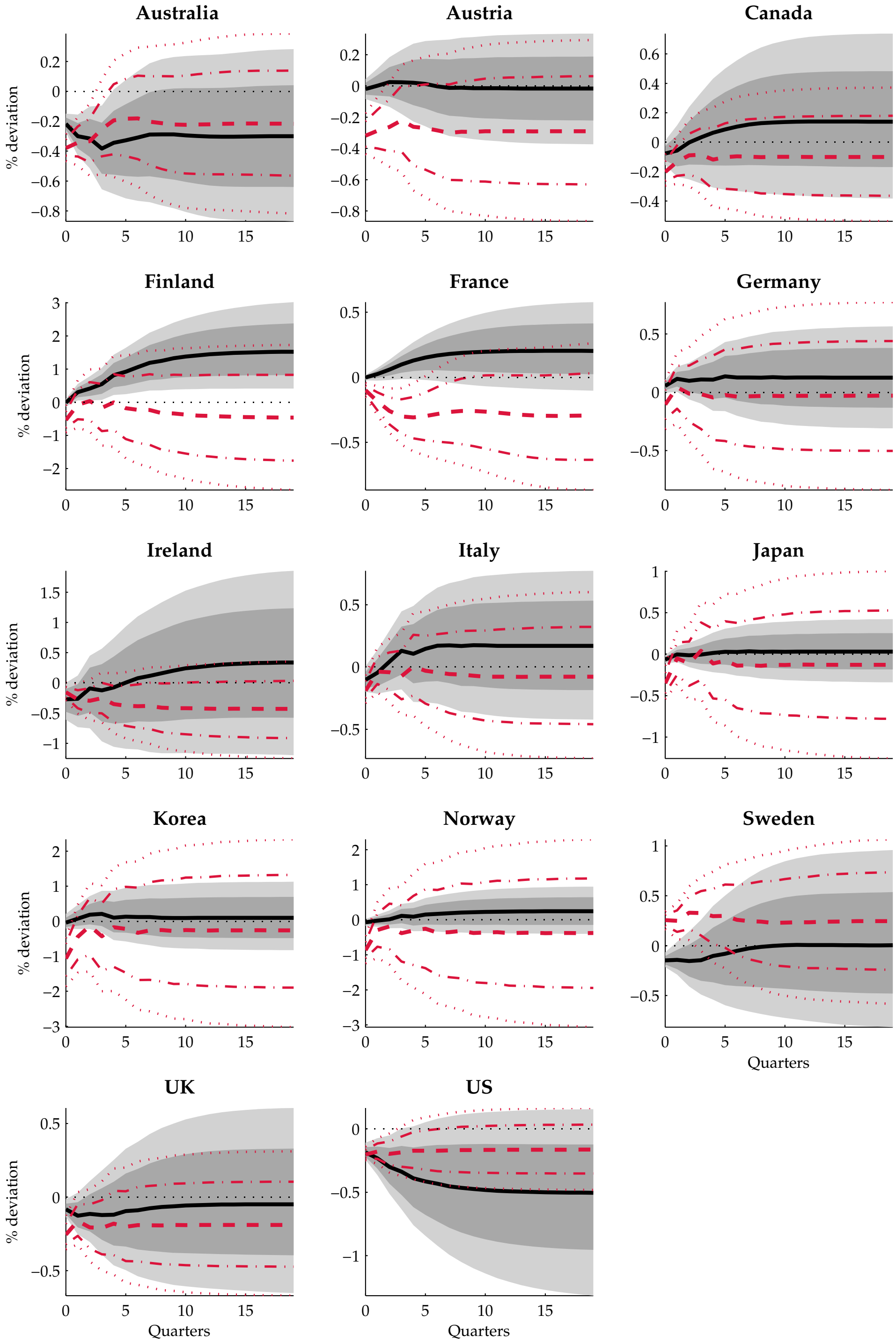
Note: Solid black lines with shaded bands depict results for our 6+1-variate baseline specification estimated via EGLS. Dashed and dotted red lines correpsond to 6-variate specification estimated via OLS. Confidence bands are obtained with 1000 bootstrap replications and cover 68% and 90%. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Figure 8: *Median fraction of FEV of aggregate hours accounted for by technology shocks*



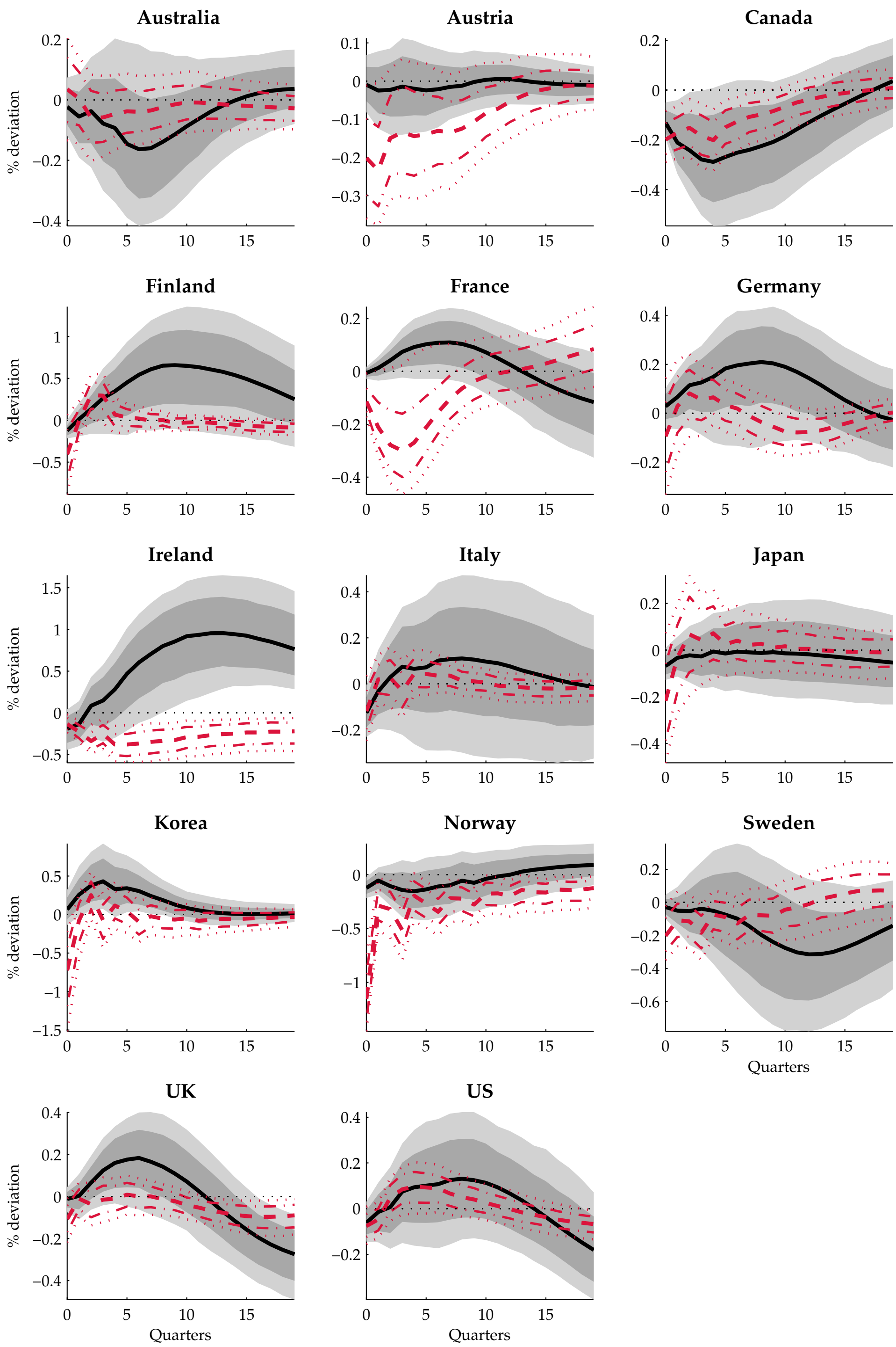
Note: Figure depicts median estimates for 6+1-variate (in black) and 2+1-variate (in red) specifications, in both cases for identification via the maxFEV approach (solid lines) and LR-restrictions (dotted lines). In the 2+1-specification, all variables enter in first-differences; in the 6+1-specification, total hours are quadratically detrended and in levels, labor productivity is adjusted for trend breaks (as reported in Table 15) and enters in first differences.

Figure 9: *Dynamic response of intensive and extensive margins to a positive technology shock in the 2+1-specification*



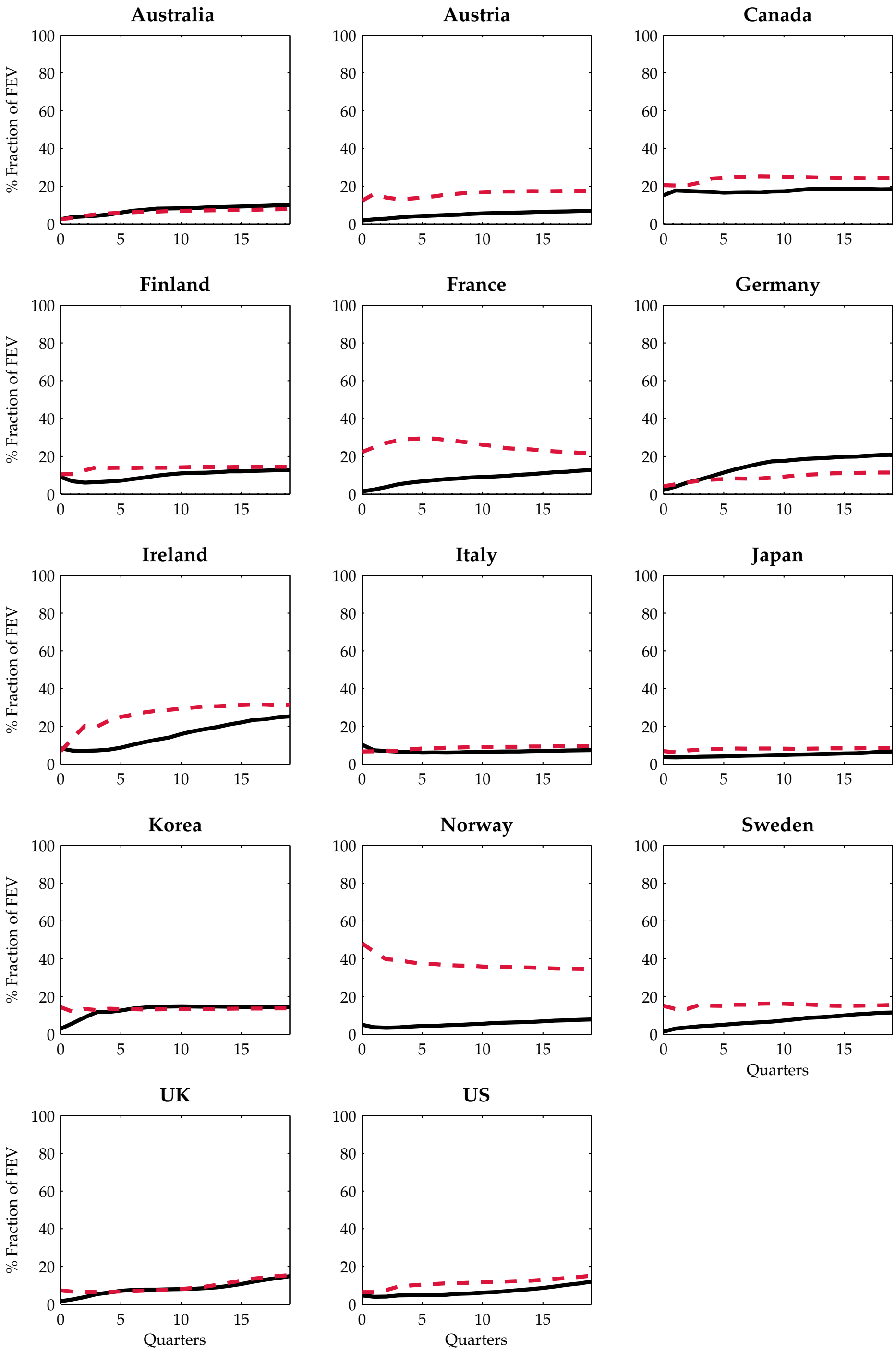
Note: The figure shows the decomposition of the IRFs of aggregate hours worked from our 2+1-specification—as depicted in dashed and dotted red lines in Figure 1—into the intensive margin (here in red) and extensive margin (here in black and gray). Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Figure 10: *Dynamic response of intensive and extensive margins to a positive technology shock in the 6+1-specification*



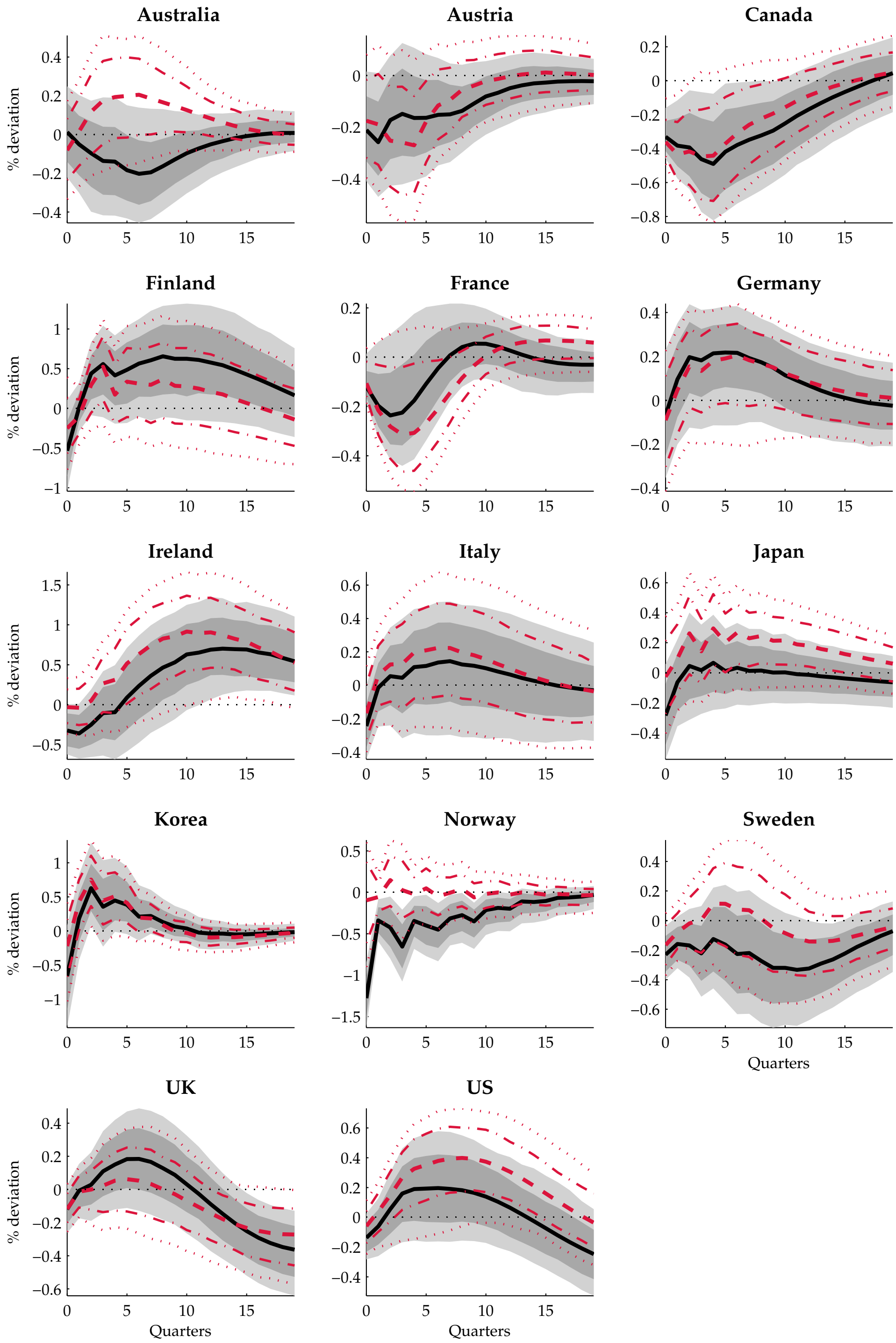
Note: The figure shows the decomposition of the IRFs of aggregate hours worked from our baseline model—as depicted in black lines and shaded bands in Figure 1—into the intensive margin (here in red) and extensive margin (here in black and gray). Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Figure 11: *Fraction of FEV of the intensive and extensive margins accounted for by technology shocks*



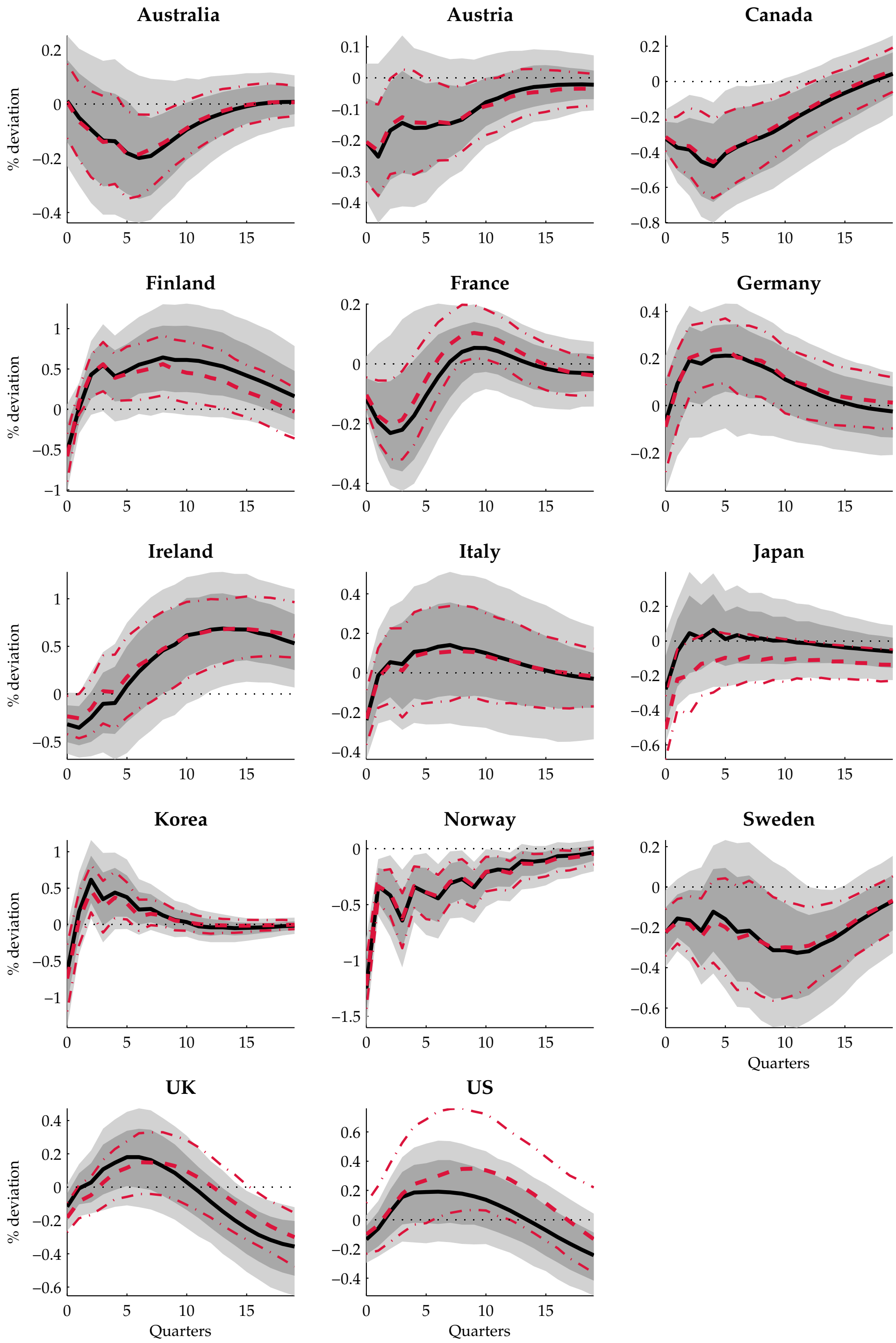
Note: The figure shows the decomposition of the FEV of total hours worked from our benchmark model—as depicted in black lines and shaded bands in Figure 8—into the intensive margin (here in dashed red) and extensive margin (here in solid black).

Figure 12: *Dynamic response of aggregate hours to a positive technology shock: Identification via maxFEV vs LR-restrictions*



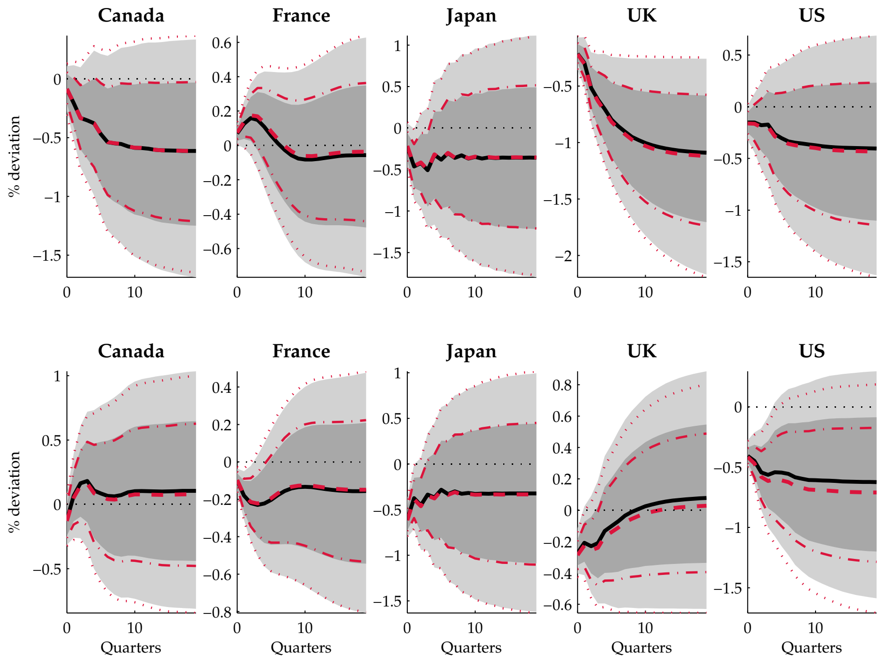
Note: Solid black lines with shaded bands depict results for the 6+1-variate baseline specification. Red lines depcit results for the same specification, but with identification via *long run restrictions* instead of the maxFEV approach. Confidence bands are obtained with 1000 bootstrap replications and cover 68% and 90%. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Figure 13: *Dynamic response of aggregate hours to a positive technology shock: GDP deflator versus CPI data*



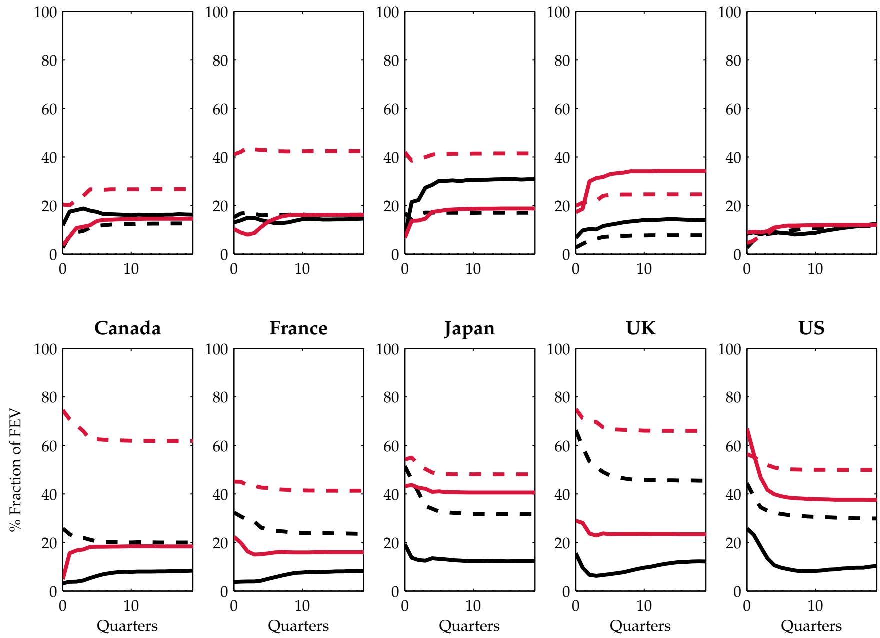
Note: Solid black lines with shaded bands depict results for the 6+1-variate baseline specification. Red lines depict results for the same specification, but with inflation measured based on CPI data, rather than the GDP deflator. Confidence bands are obtained with 1000 bootstrap replications and cover 68% and 90%. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Figure 14: *Dynamic response of total hours to positive I- (top row) and N-shocks (bottom row) in 3+1-specification*



Note: Results based on 3+1-specification with identification via maxFEV approach (solid black lines with shaded bands) and LRI (red lines). Confidence bands are obtained with 1000 bootstrap replications and cover 68% and 90%. Bias-correction of the IRFs and computation of the confidence bands has been implemented as in Kilian (1998).

Figure 15: *FEV of total hours and labor productivity attributed to I- (top row) and N-shocks (bottom row)*



Note: Figure depicts median estimates for 7+1-variate (in black) and 3+1-variate (in red) specifications for total hours (solid lines) and labor productivity (dashed lines).

Table 18: *Robustness analysis for SVAR including the relative price of investment*

	I-shock										N-shock									
Country	CA		FR		JP		UK		US		CA		FR		JP		UK		US	
Baseline	0.18	+	0.09	++	-0.33	**	0.13	+	0.13	+	-0.03		-0.03		-0.42	**	-0.18	**	-0.27	**
OLS vs EGLS	0.18	+	0.09	++	-0.32	**	0.13	+	0.13	+	-0.03		-0.03		-0.41	**	-0.18	**	-0.27	**
Lag-Length																				
1	0.30	++	0.16	++	-0.34	**	0.12		0.29	++	-0.17	*	-0.02		-0.33	**	-0.22	**	-0.17	*
2	0.16	+	0.09	++	-0.21	*	0.14	+	0.15	+	-0.17	**	-0.01		-0.47	**	-0.20	**	-0.19	*
Forecast Horizon																				
20	0.14	+	0.09	++	-0.34	**	0.08	+	0.13	++	-0.07		-0.04	*	-0.44	**	-0.13	*	-0.27	**
60	0.19	+	0.09	++	-0.31	*	0.09		0.10		-0.01		-0.01		-0.39	**	0.03		-0.28	**
Hours specification																				
Difference	0.13	+	0.09	++	-0.39	**	-0.04		0.16	++	-0.26	**	-0.12	**	-0.39	**	-0.10	*	-0.22	**
Level	0.08		0.15	+	-0.51	**	0.10	+	0.22	++	0.00		-0.02		-0.45	**	-0.01		-0.15	*
Linear	0.13	+	0.09	++	-0.50	**	0.11	+	0.23	++	0.01		-0.02		-0.46	**	-0.01		-0.15	*
Cubic	0.20	+	0.10	++	-0.05		0.13	+	0.14	+	-0.01		-0.03		-0.50	**	-0.04		-0.25	**
No trend breaks	0.18	+	0.09	++	-0.29	**	0.13	+	0.13	+	-0.03		-0.03		-0.39	**	-0.03		-0.27	**
In output units	0.06		0.11	++	-0.31	**	0.15	+	0.14	+	-0.08		-0.09	**	-0.43	**	-0.13	*	-0.27	**

Note: **/++ indicates significance at 95%-level, */+ indicates significance at 68%-level (based on 1000 bootstrap replications).