INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

Do Energy Efficiency Improvements Reduce Energy Use? Empirical Evidence on the Economy-Wide Rebound Effect in Europe and the United States

Anne Berner [a]
Stephan Bruns [b]
Alessio Moneta [c]
David I. Stern [d]

[a] Chair of Statistics, University of Göttingen, Germany.
[b] Centre for Environmental Sciences, University of Hasselt, Belgium and
Chair of Econometrics, University of Göttingen, Germany.
[c] Institute of Economics and EMbeDS, Sant'Anna School of Advanced Studies, Pisa, Italy.
[d] Arndt-Corden Department of Economics, Crawford School of Public Policy,
The Australian National University, Australia.

2021/20 May 2021
ISSN(ONLINE) 2284-0400

Do Energy Efficiency Improvements Reduce Energy Use? Empirical Evidence on the Economy-Wide Rebound Effect in Europe and the United States

Anne Berner[a], Stephan Bruns[b], Alessio Moneta[c], David I. Stern[d]

[a] *Chair of Statistics, University of Göttingen, Germany*

[b] *Centre for Environmental Sciences, University of Hasselt, Belgium and Chair of Econometrics, University of Göttingen, Germany*

[c] *Institute of Economics and EMbeDS, Sant'Anna School of Advanced Studies, Italy*

[d] *Arndt-Corden Department of Economics, Crawford School of Public Policy, The Australian National University, Australia*

Abstract

Improving energy efficiency is often considered to be one of the keys to reducing greenhouse gas emissions. However, efficiency gains also reduce the cost of energy services and may even reduce the price of energy, resulting in energy use rebounding and potential energy use savings being eaten up. There is only limited empirical research quantifying the economy-wide rebound effect that takes the dynamic economic responses to energy efficiency improvements into account. We use a Structural Factor-Augmented Vector Autoregressive model (S-FAVAR) that allows us to track how energy use changes in response to an energy efficiency improvement while accounting for a vast range of potential confounders. Our findings point to economy-wide rebound effects of 78% to 101% after two years in France, Germany, Italy, the U.K., and the U.S. These findings imply that energy efficiency innovations alone may be of limited help in reducing future energy use and emphasize the importance of tackling carbon emissions directly.

Keywords: Energy efficiency, economy-wide rebound effect, climate change, climate policy, Structural FAVAR, Independent Component Analysis

Acknowledgments: The research leading to this article has been funded by the Federal Ministry of Education and Research (BMBF) (Funding code: 01UT170). We thank the members of ReCap project team (https://www.macro-rebounds.org/) as well as participants of the IAEE Conference (2019), the ESEE Conference (2019) and seminars at the Chairs of Statistics and Econometrics in Göttingen for helpful comments.

---

1. Introduction

"Energy efficiency" is one of the key concepts of green new deal strategies to mitigate greenhouse gas emissions (IPCC, 2019; IEA, 2016). In political discussions, energy efficiency is seen as a panacea for reducing energy consumption while simultaneously reducing the costs of production and thereby ensuring green growth (European Commission, 2019; Ocasio Cortez, 2019; OECD, 2012). However, efficiency gains and the associated cost reductions will result in some rebound, whereby energy use savings due to the gains in efficiency are reduced or even completely eaten up.

The *direct* rebound effect describes the behavioral response of consumers and producers that will want to expand the use of energy services as the cost of these services falls (e.g. Sorrell and Dimitropoulos, 2008). There are also many follow-on effects across the economy known as *indirect* rebound effects. For example, a cost-saving energy efficiency gain for consumers will redirect saved income to other goods and services that also require energy to be produced. Furthermore, a cost-saving energy efficiency gain may also lower the price of energy resulting in further incentives to expand the use of energy services and the new energy-saving technology might even require more energy to be produced than the old one (Lange et al., 2020; Sorrell and Dimitropoulos, 2008; Gillingham et al., 2016). The rebound effect measures – as a percentage of the potential savings in energy use – the extent of savings in energy use that have not been realized due to the direct and indirect responses of economic agents to the initial efficiency gain.

While direct rebound effects are comparatively well studied and are on average estimated to range between 10% and 30% in developed countries (Maxwell and McAndrew, 2011), it is challenging to estimate the economy-wide rebound effect, which encompasses both direct and indirect rebound effects. In this study, we empirically estimate the economy-wide rebound effect for a sample of European countries and the United States, finding rebound effects that approach almost 100% after two years.

The quantitative literature on the economy-wide rebound effect can be divided into computational, accounting, and fully empirical approaches (Stern, 2020). Computational approaches are used most frequently, including partial equilibrium approaches (e.g. Saunders, 2008) and computable general equilibrium (CGE) models (e.g. Turner, 2009; Koesler, 2013; Rausch and Schwerin, 2018). These structural models are theoretically comprehensive and can capture a wide range of mechanisms. The estimated rebound effects from CGE models vary between negative effects, indicating that energy use is reduced by more than the efficiency improvement, to the opposite effect where energy efficiency triggers an increase in the use of energy (known as "backfire") (Turner, 2009; Colmenares et al., 2019). The accounting approach (Lin and Liu, 2012; Shao et al., 2014; Lin and Du, 2015; Zhang and Lin Lawell, 2017) treats changes in energy intensity as changes in energy efficiency and assumes that rebound is proportional to total factor productivity growth, neither of which is appropriate (Stern, 2020).

Only a few studies try to quantify the economy-wide rebound effect fully econometrically, using observed data and statistical methods (Adetutu et al., 2016; Orea et al., 2015; Yan et al., 2019). The key challenge that all these studies face when empirically estimating the effect of energy efficiency improvements on energy use is the interdependence and co-evolution of the relevant time series. Existing studies do not allow GDP and the price of energy to change in response to changes in energy efficiency. These changes in GDP and the price of energy (and also other relevant time series), however, may then result in further changes in energy use and ignoring these dependencies will bias estimates of the economy-wide rebound effect.

Recently, Bruns et al. (2021) proposed a Structural Vector Autoregressive (SVAR) model for estimating the economy-wide rebound effect. SVAR models are the workhorse of macroeconomic time series analysis consisting of a small system of regression equations that represent the statistical dependence among the relevant time series (Kilian and Lütkepohl, 2017). In this framework, exogenous changes in energy efficiency can be identified and the reaction of energy use to these changes can be measured, taking into account the possibility that this reaction may be mediated by other variables such as prices and GDP. Bruns et al. (2021) use this approach to provide a first estimate of the economy-wide rebound effect for the U.S., indicating that the rebound effect is about 100%.

In this study, we extend the work of Bruns et al. (2021) in two directions. First, while the SVAR approach provides powerful tools for estimating the responses of an economic system to exogenous forces, the presence of unobserved confounders may bias these estimates (Bernanke et al., 2005; Bai and Ng, 2013; Favero et al., 2005). Accounting for unobserved confounders in macroeconomic time series analysis is non-trivial as the number of potential confounders is very large, while the number of available observations is small. Therefore, we apply a Structural Factor-Augmented Vector Autoregressive (S-FAVAR) model. These models not only estimate, like SVAR models, the relationship among several variables over time but also augment the core model with the principal components of a rich set of potential confounders (Bernanke et al., 2005). Specifically, our core model includes three variables: energy use, the real price of energy, and GDP. We obtain the additional factors from a set of 41 to 56 economic time series depending on the country considered. This approach helps to comprehensively mitigate the threat of omitted-variable biases and to reduce the potential bias due to economic agents anticipating energy efficiency improvements (nonfundamental shocks). Second, Bruns et al. (2021) estimate a rebound effect of roughly 100% for the U.S., but it is also important to investigate whether energy efficiency innovation in large polluting countries other than the U.S. is equally unlikely to significantly reduce energy use in the long run. Therefore, we use the improved S-FAVAR approach to estimate economy-wide rebound effects in France, Germany, Italy, the U.K., and the U.S.

Our analysis relies on the notion that changes in the economic system can be traced back to independent impulses, commonly referred to as "shocks" in the econometrics literature (Kilian and Lütkepohl, 2017). We identify an energy efficiency shock by applying Independent Component Analysis (ICA) to the residuals of a reduced-form Factor-Augmented Vector Autoregressive (FAVAR) model. ICA finds the least dependent linear combinations of the residuals, which correspond to an estimate of the independent shocks that jointly affect the observed variables. Based on this, we can estimate the response of the economy-wide energy use over time to an energy efficiency shock.

We find that the economy-wide rebound effect ranges from 78% to 101% for France, Germany, Italy, the U.K., and the U.S. after two years. This implies that policies to encourage energy efficiency improvements may not be effective in reducing energy use in the long run, which might be at odds with common green growth strategies.

The remainder of the paper is organized as follows. Section 2 presents the emprirical strategy that is used to estimate economy-wide rebound effects, disentangling the different components of the S-FAVAR model and introducing the dataset. Empirical results are discussed in Section 3. Finally, Section 4 summarizes and concludes.

2. Empirical Approach

2.1. The Economy-Wide Rebound Effect

The economy-wide rebound effect is defined as the extent of savings in energy use across the economy that have not been realized due to the direct and indirect responses of economic agents to the initial efficiency gain. We estimate the economy-wide rebound effect by identifying an energy efficiency shock, that is, an independent and exogenous shock to economy-wide energy use that cannot be explained by any other variable considered in the S-FAVAR model as outlined in the subsequent sections, and by tracing the dynamic response of energy use to this shock. Using the subscript i to denote the number of periods since the energy efficiency improvement, the economy-wide rebound effect is given by:

$$R_i = 1 - \frac{\text{Actual}}{\text{Potential}} = 1 - \frac{\Delta \hat{e}_i}{\varepsilon_{e_1}} \tag{1}$$

where ε_{e_1} is the contemporaneous response of energy use to the energy efficiency shock, which represents the potential "engineering" change in energy use, and $\Delta \hat{e}_i$ is the actual change in energy use (Bruns et al., 2021). Notice that ε_{e_1} is by construction a negative number, while $\Delta \hat{e}_i$ measures the response of energy use to the energy efficiency shock after i periods and can be any real number.

2.2. Structural Factor-Augmented Vector Autoregressive (S-FAVAR) model

It would be desirable to consider all variables that potentially influence economy-wide energy use and, therefore, potentially confound the estimate of the economy-wide rebound effect. However, the analysis of intertemporal dependencies in a "data-rich" environment is problematic using standard multivariate autoregression models, as the number of parameters to be estimated may rapidly exceed the available observations. Augmenting a classical SVAR model with factors obtained from a large set of time series provides a remedy.

To characterize the effect of an efficiency shock on energy use, we assume that the state of the economy is represented by a vector C_t, whose entries are both observed and latent variables. As we are interested in estimating the response of energy use to an energy efficiency shock, we include the following three observable series: Energy use (E_t), GDP (Y_t), and the price of energy (P_t). We consider these three variables to be the core variables when analyzing economy-wide rebound effects. Moreover, we incorporate several latent factors (F_t) in the vector C_t that summarize the information in a large set of macroeconomic indicators (see Section 2.3 for the estimation of these factors). The dynamics of the common components are modeled by the following reduced-form FAVAR model

$$C_t = \Phi(L)C_{t-1} + u_t$$

where

$$C_t = \begin{bmatrix} E_t \\ Y_t \\ P_t \\ F_t \end{bmatrix} \tag{2}$$

where $\phi(L)$ is a conformable lag polynomial of finite order and the error term and u_t is assumed to be i.i.d. with mean zero.

In contrast to a traditional VAR model, the system in Equation (2) also includes latent factors, F_t. These factors can be extracted from a large number of macroeconomic time series. The dynamic factor model is explained in the following subsection.

2.3. Factor augmentation

The general idea of the factor model is to reduce a large matrix of time series data into a few latent factors. The following equation relates the unobserved common factors, collected in the $r \times 1$ vector F_t, and the vector of m observed core variables W_t (in our case time series data on the price of energy, energy use and GDP, so that $m = 3$) to an $N \times 1$ vector of (observed) "informational" variables Z_t (in our case 41 to 56 time series, depending on the country analyzed):

$$Z_t = \Lambda^f F_t + DW_t + \zeta_t, \tag{3}$$

where Λ^f is an $N \times r$ matrix of factor loadings, D is a $N \times m$ diagonal matrix, and ζ_t is a $N \times 1$ vector of idiosyncratic residuals. Hence, changes in Z_t are driven by the latent factors (F_t) and the endogenous observable time series (W_t), plus idiosyncratic noise. In the multi-period setting $Z = (Z_1, Z_2, \dots, Z_T)'$ is the $T \times N$ data matrix, $W = (W_1, W_2, \dots, W_T)'$ a $T \times m$ matrix of observables and $F = (F_1, F_2, \dots, F_T)'$ is a $T \times r$ matrix of latent factors. We follow Hwang (2009) to obtain estimates of D, Λ^f and F by the following steps:

1. Regress Z_t on W_t, and compute the least squares estimates, $\hat{D}$, and the residuals $\hat{U}_t = Z_t - \hat{D}W_t$, with $\hat{U} = (\hat{U}_1, \ldots, \hat{U}_T)'$;
2. Estimate the first $K - r$ principal components of $\hat{U}_t$ which represent the estimated latent factors.

Hence, the factor estimates can be specified as $\hat{F} = \hat{U}'\Lambda^f$, where the columns of Λ^f are the eigenvectors corresponding to the largest eigenvalues of $\hat{U}'\hat{U}$. This ensures that the loading matrix has orthonormal columns and can be identified.[1]. The resulting factors, F_t, are included in the reduced-form FAVAR in Eq. (2), which can be estimated using OLS, before identifying the structural representation.

2.4. Identification

After estimating the factors, the model in Equation 2 can be treated as a standard VAR. As the residuals, u_t, in equation (2), might be correlated across equations, we rewrite these innovations as a linear combination of the underlying orthogonal structural disturbances η_t. Rewriting Equation (2) results in the following structural model

$$\begin{bmatrix} F_t \\ W_t \end{bmatrix} = \phi(L) \begin{bmatrix} F_{t-1} \\ W_{t-1} \end{bmatrix} + B\eta_t \tag{4}$$

where η_t, has mean zero with covariance matrix Σ. The non-singular matrix B specifies the contemporaneous relations between the shocks and the reduced-form innovations $u_t{=}B\eta_t$, with $E[u_t] = 0$ and $\text{Cov}[u_t] = BB' = \Sigma_u$. The mixing matrix, B, contemporaneously transmits the effects of the shocks to the dependent variables.

The matrix B is estimated and hence the shocks are identified using two different search methods that use unsupervised statistical learning typical of machine learning research and fall under the class of Independent Component Analysis (Comon, 1994). Both methods rely on two key assumptions about the statistical properties of the vector of shocks. Namely, the shocks are assumed to be mutually statistically independent and distributed according to a (not necessarily specified) non-

[1]See Kilian and Lütkepohl (2017) Table 16.1 or Bai and Ng (2013) for alternative sets of identification conditions for factors and factor loadings.

Gaussian distribution, with at most one exception. The latter assumption can be easily checked indirectly by testing whether Gaussianity of the reduced-form innovations, u_t, is rejected. The former assumption cannot be tested, but is in tune with the idea of finding the primitive exogenous forces that drive the dynamics of the system, each of which is denoted by a particular economic characteristic, not shared with the others, and that can be possibly used as policy levers.

The two ICA approaches we apply are distance covariance (dcov) (Matteson and Tsay, 2011) and non-Gaussian Maximum Likelihood (ngml) (Lanne et al., 2017), which have been recently studied in the econometric literature in the context of SVAR models (Herwartz, 2018). We further probe the robustness of our results by computing the Choleski decomposition of the residual variance matrix which gives similar results (see Table D.6 for a comparison of the different rebound estimations).

ICA does not determine the sign nor the economic meaning of the shocks *a priori*. The columns of the (instantaneous) impact matrix should be reordered and if necessary their sign changed to make them easier to interpret economically (Gouriéroux et al., 2017; Moneta and Pallante, 2020).[2] We solve this indeterminacy by assuming that of the three empirically identified shocks the energy efficiency improvement should have the largest (in absolute value) contemporaneous effect on energy use. This shock represents exogenous changes to energy use that are not explained by any of the other variables considered in the model and, thus, we attribute this exogenous change to a change in efficiency. The effect of this shock on energy use is by definition negative as we are interested in studying the effect of improvements on energy efficiency.

In our analysis, we extensively use the R package `svars`, which implements independence-based identification (Lange et al., 2019).

2.5. Estimating the economy-wide rebound effect

The rebound effect is defined as the ratio between actual and potential energy savings (see equation (1)), which can be approximated by the evolution of the impulse-response functions. Figure 1 shows an illustrative impulse-response function of energy use with respect to an energy-specific shock. Here the initial or potential savings (ε_{e_1}), indicated at time 0, decrease over time and even exceed, in

[2] In the language of matrix analysis, ICA identifies the impact matrix up to the right multiplication of a signed permutation matrix, i.e. a matrix containing exactly one entry in each row and column equal to +1 or -1 and all other entries equal to 0. ICA leaves undetermined also the scale of the shocks, but these are typically normalized to have unit variance.

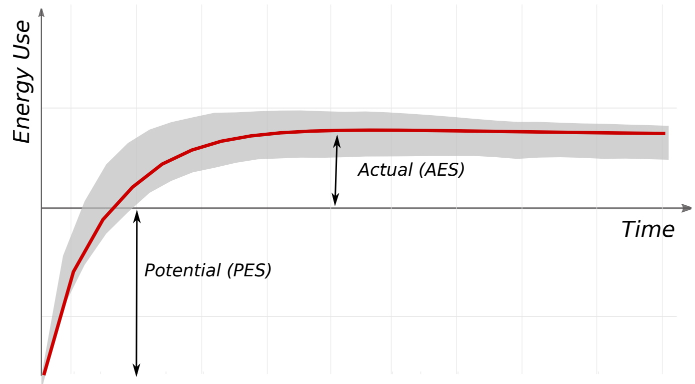

Figure 1: **Illustration of potential energy savings (PES) and actual energy savings (AES)** depicting an exemplary impulse-response function of energy use (red curve) with an exemplary confidence interval (gray area).

this particular illustration, the pre-shock level leading to actual savings ($\Delta\hat{\varepsilon}_i$) that are negative and, therefore, to backfire.

The estimation of the rebound effect based on an S-FAVAR model addresses the omitted variables problem that is common in SVAR analysis by including the information from a large set of variables. Furthermore, the S-FAVAR model allows us to tackle a related but subtler problem, which is typical of standard (small scale) SVAR models and may bias the estimation of the rebound effect. In SVAR analysis, structural shocks are identified from a linear transformation of VAR prediction errors (i.e. reduced-form residuals). But it is conceivable that these prediction errors do not accurately capture the true prediction errors of the economic agents, because the latter rely on an information set that is larger than the one contained in the econometric model. This creates a mismatch between the shocks of the (true) data generating process and the shocks of the SVAR model, which has been studied in the literature on so-called *nonfundamental shocks*[3] (Kilian and Lütkepohl, 2017; Alessi et al., 2011). In case of such a mismatch, the shocks identified from a VAR model may in fact be anticipated by economic agents. This would engender a bias in the estimate of the energy efficiency shock and of its rebound effect. This problem, and, more generally, the problem of nonfundamental shocks, can be resolved or at least ameliorated in S-FAVAR analysis because the information set is enlarged and so it is more likely that it mirrors the information set that economic agents use to

[3]The name is due to the fact that the moving average representation of the VAR prediction errors is called the fundamental representation. Nonfundamental shocks are shocks that cannot be recovered from this representation.

predict or anticipate energy efficiency innovations.

However, our estimation still bears two caveats that need to be considered. First, the model does not capture rebound that may happen contemporaneously with the efficiency improvement. This effect is discussed in the literature under the terms "embodied-energy effect" or "redesign effect" (Lange et al., 2021). Bruns et al. (2021), however, explain that this error is smaller the closer the true rebound effect is to 100%. Second, our rebound-measure describes only the response that can be attributed to energy-specific efficiency improvements. The reason is that our energy efficiency shock is assumed to be orthogonal to other shocks. Hence, if labor- or capital-augmenting innovations are captured in the GDP shock (or other shocks) and if these innovations are correlated with improvements in energy efficiency, then these energy efficiency improvements are not captured in the energy efficiency shock.

2.6. Data

The main variables of our model comprise energy use, the price of energy, and economic output, measured by GDP. For the U.S., the data used in this article corresponds to the data described in Bruns et al. (2021). Compared to the U.S., in Europe monthly time series data at the country level are still quite sparse. Therefore, we restrict our analysis to France, Germany, Italy, and the U.K. as the monthly data for these countries and variables are available from January 2008 to September 2019, providing 141 observations. All data series were log-transformed and deseasonalized using the `seasonal` package in R with the X-11 adjustment procedure.

Additionally, the extraction of the latent factors is based on a large matrix of time series describing the economy. For this purpose, we use the Main Economic Indicator (MEI) database which is developed and maintained by the OECD. This data set presents comparative statistics that provide an overview of recent international economic developments for the European countries we analyzed, covering information on the labor market, national accounts, retail sales, production, construction, prices, finance, international trade, and the balance of payments (OECD, 2018). The latent factors are intended to summarize the main source of variation in the data panel and hence can be interpreted as common driving forces behind different variables of the economy. Online Appendix A discusses the sources of the data in detail.

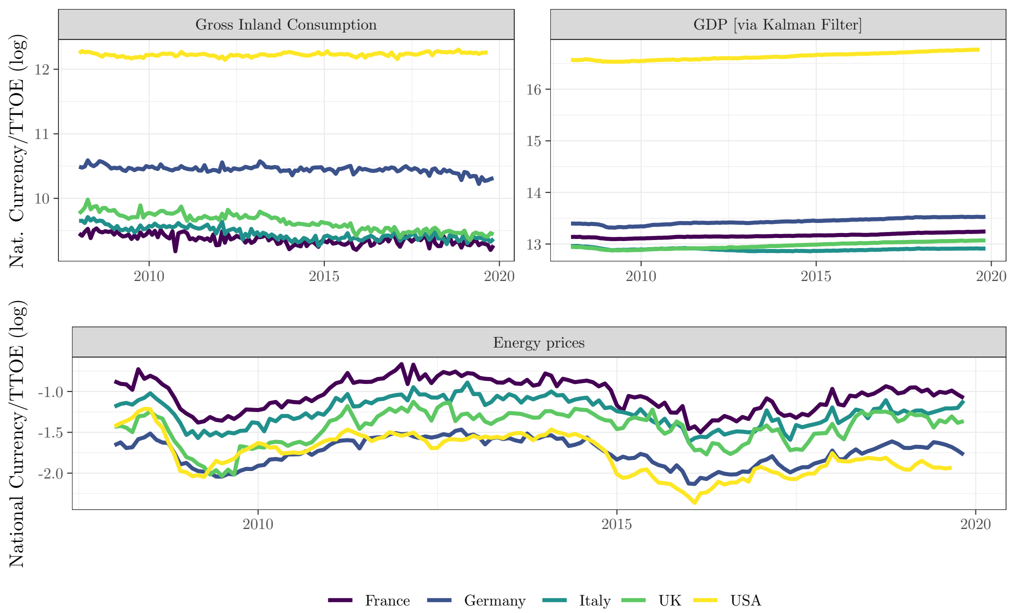

Figure 2: **Time series data for the countries included in the analysis.**

3. Results

3.1. Reduced-form FAVAR

Using the Akaike information criterion, we select lag lengths of $p = 2$ for France, the U.K. and the U.S., $p = 3$ for Italy and $p = 4$ for Germany. Maximum lag lengths of 6 and 12 both result in selecting the same lag length (see Table D.5 in the Online Appendix for the comparison).

We statistically evaluate the number of Gaussian components based on component-wise normality tests (Shapiro-Wilk, Shapiro-Francia, Jarque-Bera) for the reduced-form residuals of the model.[4] The test results indicate that the presence of more than one Gaussian component cannot be rejected (see Table D.7 in the Online Appendix). However, these tests perform poorly in small samples, which is particularly true if the distributions of the samples are close to normality (Gouriéroux et al., 2017; Maxand, 2018). Maxand (2018) show that at least the unique identification of the non-Gaussian shocks can be guaranteed irrespective of the distributions of the remaining system. We are

[4]Additionally, we compared the component-wise tests with a boostrappping test, based on forth order blind identification (FOBI) explained in the Online appendix

especially interested in the energy efficiency shock and at least the normality of the reduced-form residuals of the energy use equation can be rejected for all countries except France. Furthermore, in the case of multiple Gaussian reduced-form residuals, the ICA methods in any case deliver orthogonal shocks, since they orthogonalize the residuals like a standard principal component analysis. However, the residuals are identified up to an orthogonal transformation, which may dramatically increase the variance of the estimates (Hyvärinen and Oja, 2000). Additionally, we tested the robustness of the identified shocks by comparing the result of the independence-based identification strategies with the results derived by a Choleski decomposition finding similar results for the energy efficiency shock (see D.6).

3.2. Factor augmentation to account for potential confounders

The first two factors explain from 45.78% (U.K.) to 62.82% (U.S.) of the variance of the informational variables in each country panel (see Table 1). We include these two factors in the S-FAVAR model to ensure a balance between the variance explained and degrees of freedom considerations. Increasing the number of included factors by one adds roughly 10% to the explained variance (see Table 1). A robustness check of the estimated rebound effect with three factors included can be found in the Online Appendix (Figure D.17).

Table 1: Explained variance in the set of country-specific time series

Factor	1	2	3	4	5	6	7	8	9	10
France	33.22	13.49	11.31	7.69	6.46	6.17	5.56	5.07	4.82	4.08
Germany	35.59	19.54	12.09	8.89	8.39	5.73	4.36	4.17	4.06	3.60
Italy	37.34	15.04	10.27	8.80	6.97	6.06	5.65	5.43	4.21	3.64
UK	23.78	22.00	11.70	9.88	7.23	6.11	5.81	5.32	4.55	4.11
USA	43.33	19.49	12.80	9.16	8.29	6.57	5.05	4.41	3.72	3.44

Notes: Each row shows the variance in the country-specific set of time series explained by the respective factor (in %).

These two estimated latent factors are presented in Figure 3. The identification of the estimated factors is only possible up to a change of sign.[5] The factors fluctuate strongly during the financial

[5]This becomes evident as Factor 1 peaks during the financial crisis in 2008/2009 for Germany, Italy and the U.K. and collapses in France and the U.S.

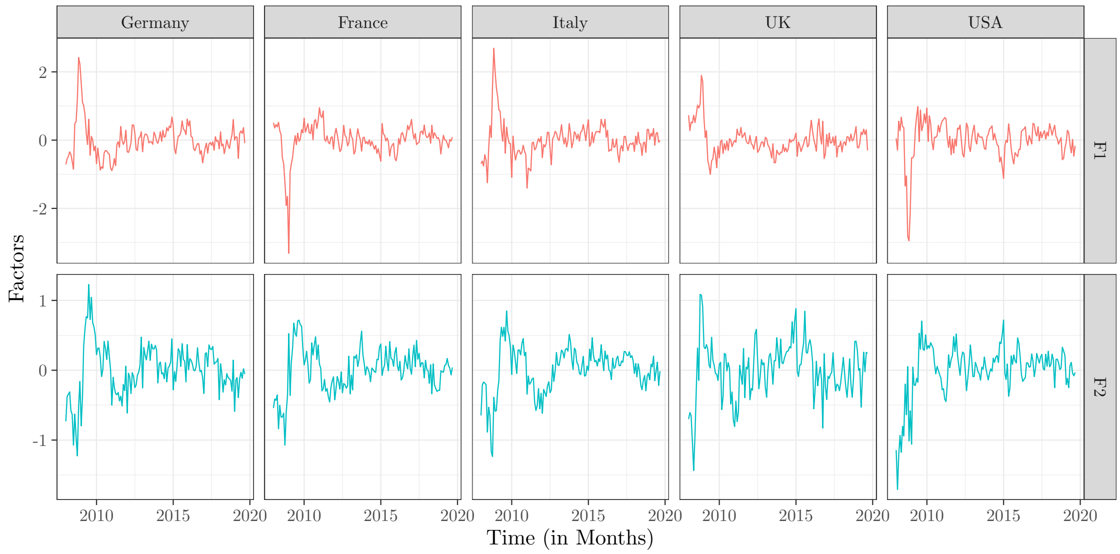

Figure 3: **Estimated latent factors.** The factors with the highest explanatory power, factor 1 (in red) and factor 2 (in blue), are depicted for each country.

crisis that started in 2008, which indicates that they enlarge the information set by adding the impact of the financial crisis.

We present the factor loadings for Germany (Panel a) and the U.K. (Panel b) in Figure 4 to show what the latent factors might represent. The higher the absolute value of the factor loading, the higher the correlation between the time series and the respective factor. For both countries one factor seems to load mainly on different producer price indicators and the other on exchange rates, the unemployment level, exports, industrial production, and expectations. This means that one factor mostly represents real changes in the economy while the other mostly represents changes in prices. The factor loadings for the other countries are similar to the German example and can be found in the Online Appendix (Figure D.16).

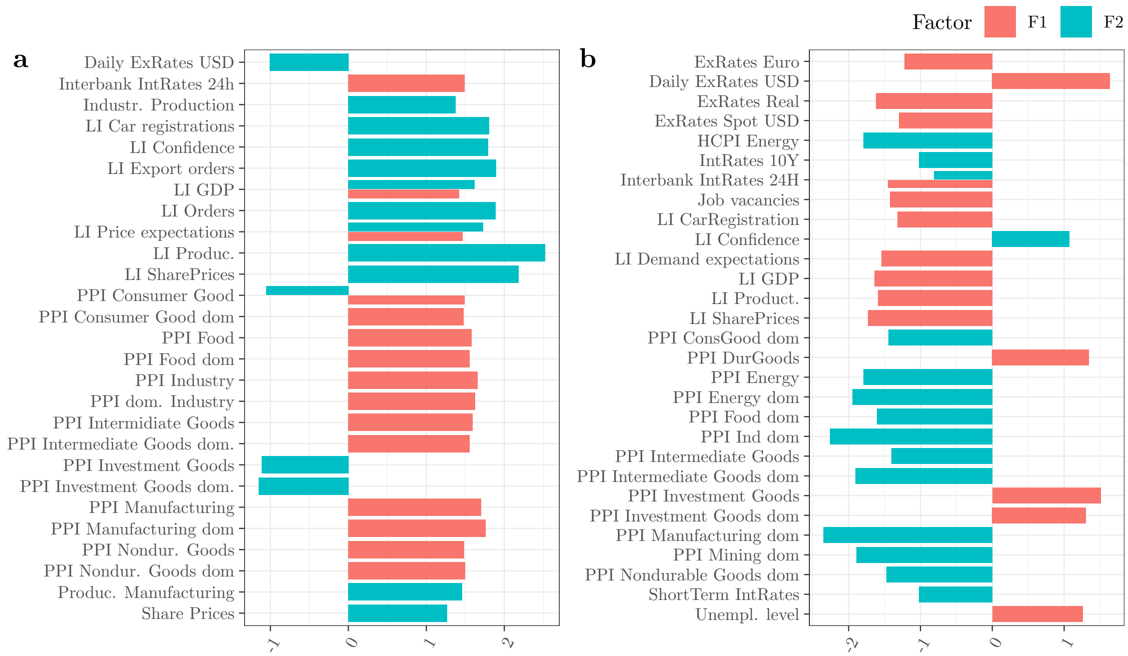

Figure 4: **Example factor loadings.** The 15 highest factor loadings for the first two factors for (**a**) Germany and the (**b**) U.K.. “ExRates” stands for exchange rates, “IntRates” for interest rates, “LI” for leading indicator, “PPI” for producer price index and “HCPI” for harmonized consumer price index.

3.3. Identifying energy efficiency shocks

We estimate a S-FAVAR model with energy use, GDP, the price of energy, and the two factors estimated in the previous stage. As described in the methods section, we identify the energy efficiency shock by using the criterion that this shock should have the largest contemporaneous effect on energy use. As our focus is on estimating the economy-wide rebound effect, identification of the energy efficiency shock is sufficient. The shocks associated with GDP and the price of energy, as well as the overall economic plausibility of the estimated S-FAVAR model, are discussed in Online Appendix C.

The identified energy efficiency shocks are presented in Table 2. For all countries, the energy efficiency shock has a large contemporaneous effect on energy use compared to its effects on GDP and the price of energy, except for the U.S. where its effect on energy use is similar in magnitude to its effect on the price of energy. The effect of this shock on energy use is negative by construction and in all countries the confidence intervals do not overlap zero. By contrast, the confidence intervals of

Table 2: Contemporaneous effects of the energy efficiency shock

	Germany	France	Italy	U.K.	U.S.
e_t	-3.41	-5.25	-4.53	-4.02	-1.8
	(-3.7, -1.32)	(-6.05, -1.85)	(-4.61, -3.51)	(-4.51, -1.46)	(-2.01, -0.68)
y_t	0.03	-0.17	-0.03	0.03	-0.01
	(-0.16, 0.14)	(-0.23, -0.02)	(-0.13, 0.08)	(-0.11, 0.11)	(-0.22, 0.17)
p_t	1.47	3.88	-1.28	-0.23	1.76
	(-0.92, 3.4)	(-0.48, 5.88)	(-3.08, 2.04)	(-3.43, 3.58)	(-0.91, 3.43)

Notes: Contemporaneous effects of the energy efficiency shock on energy use (e_t), GDP (y_t) and the price of energy (p_t) for the five countries. 95% confidence intervals in parentheses using a wild bootstrap.

the contemporaneous effects of the energy efficiency shock on GDP and the price of energy always overlap zero, except for the effect on GDP in France where zero is marginally excluded.

We corroborate the identification of the energy efficiency shock by inspecting the forecast error variance decompositions (FEVD) shown in Figure 5 (Uhlig, 2005; Netšunajev and Glass, 2017). FEVDs are a measure of the impacts of the shocks on each of the modeled variables. FEVDs show how much of the variance of the forecasting error of each variable (the prediction mean squared error of the model variables) at different time horizons is accounted for by the different shocks. If a shock accounts for most of the forecast error variance of a specific variable x, at most time horizons, this provides good evidence that the shock should be labeled as the x-shock.

The panels show for each country the percentage of the forecast error variance of energy use explained by the different shocks in the months following a shock of each type. If the forecast error variance of energy use can be largely explained by the shock that we identified as the energy efficiency shock, then this would be a strong sign that the identification is correct. For all forecast horizons in Germany, for example, about 75% of the forecast error variance of energy use is explained by the shock that we identified as the energy efficiency shock (top left plot in Figure 5). For all countries and at all time steps considered, the forecast error variance of energy use is mostly explained by the identified energy efficiency shocks.

The FEVDs for the other variables, shown in Figure B.9 of the Online Appendix, and the discussion of the economic plausibility of the estimated impulse response functions (provided in Appendix C)

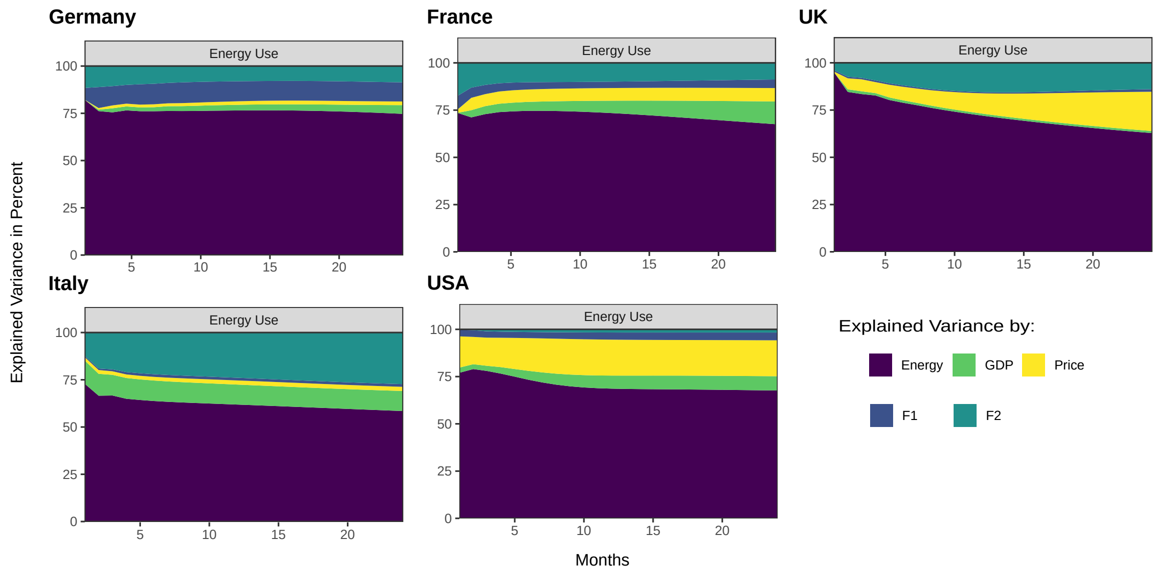

Figure 5: **Forecast error variance decomposition for energy use.** The decomposition shows the percentage (y axis) of the i-months (x axis) ahead forecast error variance which is explained by the five different shocks (indicated by the five different colours).

further strengthen our identification of the energy efficiency shock.

3.4. Economy-wide rebound effect

In Figure 6 (left panel) the impulse response functions of energy use for an energy-efficiency shock show the same tendency for all countries: after an immediate reduction in energy use due to increased efficiency, energy use rebounds towards the original level of consumption. The impulse-response curves of the U.S. and France seem to rebound faster than the other countries. However, the differences are subtle and the confidence intervals are overlapping. Figure 6 (right panel) shows that after 24 months the estimated rebound effect ranges between 78% and 101% for all countries with all confidence intervals overlapping 100%.

In general, estimates for the rebound effects tend to be consistent across countries and identification methods (compare Table D.6 in the Online Appendix).

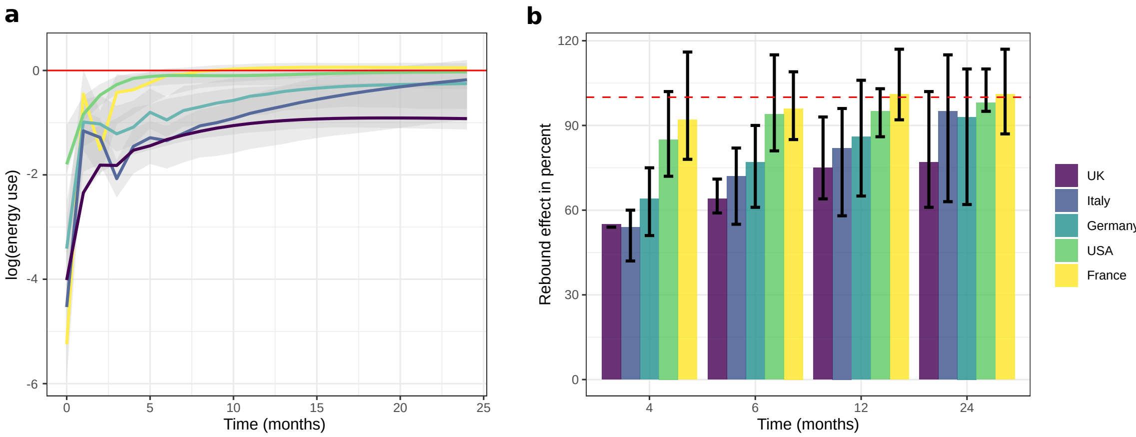

Figure 6: **Impulse response functions of energy use for an energy efficiency shock (a) and estimated rebound effects (b)**. Shaded areas represent 90% confidence intervals in the left panel. Error bars represent 90% confidence intervals in the right panel. Confidence intervals based on wild bootstrapping.

4. Discussion and Conclusions

We used a Structural Factor Augmented Vector Autoregressive (S-FAVAR) model to quantify the economy-wide effect of energy efficiency improvements on energy use. Our methodology improves on past research by being able to separate the effect of energy efficiency improvements on energy use from other factors that might influence energy use, such as economic growth, changes in the price of energy, and a multitude of other potentially confounding factors by incorporating a large number of economic time series into the analysis. Our approach also allows GDP and the price of energy to evolve in response to the energy efficiency impulse and, in turn, energy use to respond again to the evolution of GDP and the price of energy.

Our analysis extends in two main ways the work of Bruns et al. (2021) who use U.S. data to provide the first SVAR-based quantification of the economy-wide rebound effect. First, we augment the SVAR with factors obtained from a rich panel of time series to address the potentially large number of confounders. Addressing potential omitted-variable biases is crucial to improving and ensuring the reliability of the estimated economy-wide rebound effect. At the same time, augmenting the model with factors from a rich macroeconomic data set improves the information contained in the model and better reflects that available to economic agents in the real world. In this manner, it is less likely, in comparison with a small scale econometric model (e.g. SVAR), that the identified energy efficiency shocks are events that can be systematically anticipated by economic agents, which would bias the estimate of the economy-wide rebound effect. Rather, the shocks can be interpreted as genuine innovations, whose rebound effect can be reliably estimated. Second, we apply the improved estimation approach to both the U.S. and a set of European countries (France, Italy, Germany and U.K.) to explore how similar the economy-wide rebound effect is across large, high income countries.

We find that the economy-wide rebound effect is close to 100 % across our sample of countries, supporting the findings of Bruns et al. (2021). This implies that energy efficiency improvements that save energy due to the adoption of more efficient cost-reducing technology will have limited long-run impact on aggregate energy consumption.

Our analysis identifies exogenous changes in energy use as changes in energy efficiency, as they can be neither explained by the core variables nor by the additional factors. We interpret these exogenous changes to largely represent cost-reducing improvements in energy efficiency. It should be emphasized that Fullerton and Ta (2020) show in a theoretical model that energy efficiency mandates that raise the cost of energy services can have a negative rebound effect resulting in more energy being saved than mandated. On the other hand, they find that cost-reducing innovations in the face of binding energy efficiency mandates are expected to have an especially large rebound effect.

We conclude by emphasizing that even though cost-reducing energy efficiency innovations might enhance welfare, by providing more energy services to consumers and producers for a given cost, the magnitude of the estimated rebound-effect means that they will not significantly reduce energy use in the long run. However, a tightening cap on carbon emissions or equivalent carbon tax policy would reduce fossil fuel use regardless of the rebound effect. In fact, improving energy efficiency would help reduce the welfare cost of such a policy.

Supplementary material

The supplementary material contains the Online Appendix as well as data and code to reproduce all findings reported in this article.

References

Adetutu, M. O., Glass, A. J., and Weyman-Jones, T. G. (2016). Economy-wide estimates of rebound effects: Evidence from panel data. *The Energy Journal*, 37(3).

Alessi, L., Barigozzi, M., and Capasso, M. (2011). Non-fundamentalness in structural econometric models: A review. *International Statistical Review*, 79(1):16–47.

Bai, J. and Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, 176(1):18–29.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120(1):387–422.

Bernanke, B. S., Gertler, M., Watson, M., Sims, C. A., and Friedman, B. M. (1997). Systematic monetary policy and the effects of oil price shocks. *Brookings papers on economic activity*, 1997(1):91–157.

Bruns, S. B., Moneta, A., and Stern, D. I. (2021). Estimating the economy-wide rebound effect using empirically identified structural vector autoregressions. *Energy Economics*, page 105158.

Chow, G. C. and Lin, A.-l. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The review of Economics and Statistics*, pages 372–375.

Colmenares, G., Löschel, A., and Madlener, R. (2019). The Rebound Effect and its Representation in Energy and Climate Models. *SSRN Electronic Journal*, (16).

Comon, P. (1994). Independent component analysis, a new concept? *Signal processing*, 36(3):287–314.

European Commission (2019). The European Green Deal. Technical Report 9, Brussels.

Favero, C. A., Marcellino, M., and Neglia, F. (2005). Principal components at work: The empirical analysis of monetary policy with large data sets. *Journal of Applied Econometrics*, 20(5):603–620.

Fernandez, R. B. (1981). A methodological note on the estimation of time series. *The Review of Economics and Statistics*, 63(3):471–476.

Fullerton, D. and Ta, C. L. (2020). Costs of energy efficiency mandates can reverse the sign of rebound. *Journal of public economics*, 188:104225.

Gillingham, K., Rapson, D., and Wagner, G. (2016). The Rebound Effect and Energy Efficiency Policy. *Review of Environmental Economics and Policy*, 10(1):68–88.

Gouriéroux, C., Monfort, A., and Renne, J.-P. (2017). Statistical inference for independent component analysis: Application to structural VAR models. *Journal of Econometrics*, 196(1):111–126.

Herwartz, H. (2018). Hodges–Lehmann detection of structural shocks–an analysis of macroeconomic dynamics in the euro area. *Oxford Bulletin of Economics and Statistics*, 80(4):736–754.

Hwang, H.-s. (2009). Two-step estimation of a factor model in the presence of observable factors. *Economics Letters*, 105(3):247–249.

Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural Networks*, 13(4-5):411–430.

IEA (2016). World Energy Outlook 2016. Technical report, IEA, Paris.

IPCC (2019). Summary for Policymakers. In P.R. Shukla, J. Skea, E. Calvo Buendia, V. Masson-Delmotte, H.- O. Pörtner, D. C. Roberts, P. Zhai, R. Slade, S. Connors, R. van Diemen, M. Ferrat, E. Haughey, S. Luz, S. Neogi, M. Pathak, J. Petzold, J. Portugal Pereira, P. Vyas, E. Huntley, K. Kissick, J. M., editor, *Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.*

Issler, J. V. and Notini, H. H. (2016). Estimating brazilian monthly gdp: A state-space approach. *Revista Brasileira de Economia*, 70(1):41–59.

Kilian, L. and Lütkepohl, H. (2017). *Structural vector autoregressive analysis*. Cambridge University Press.

Koesler, S. (2013). Catching the rebound : economy-wide implications of an efficiency shock in the provision of transport services by households. *ZEW discussion papers*.

Lange, A., Dalheimer, B., Herwartz, H., and Maxand, S. (2019). svars: An r package for data-driven identification in multivariate time series analysis. *Journal of Statistical Software*.

Lange, S., Kern, F., Peuckert, J., and Santarius, T. (2020). A multi-level typology: Making sense of the diversity of rebound effects. *Energy Research & Social Science (in review).*

Lange, S., Kern, F., Peuckert, J., and Santarius, T. (2021). The Jevons paradox unravelled: A multi-level typology of rebound effects and mechanisms. *Energy Research and Social Science*, 74:101982.

Lanne, M., Meitz, M., and Saikkonen, P. (2017). Identification and estimation of non-Gaussian structural vector autoregressions. *Journal of Econometrics*, 196(2):288–304.

Lin, B. and Du, K. (2015). Measuring energy rebound effect in the Chinese economy: An economic accounting approach. *Energy Economics*, 50:96–104.

Lin, B. and Liu, X. (2012). Dilemma between economic development and energy conservation: Energy rebound effect in china. *Energy*, 45(1):867–873.

Matteson, D. S. and Tsay, R. S. (2011). Dynamic orthogonal components for multivariate time series. *Journal of the American Statistical Association*, 106(496):1450–1463.

Maxand, S. (2018). Identification of independent structural shocks in the presence of multiple gaussian components. *Econometrics and Statistics*.

Maxwell, D. and McAndrew, L. (2011). Addressing the Rebound Effect (Final Report No. ENV.G.4/FRA/2008/0112). Technical report, European Commission DG ENV, Ivry-Sur-Seine, France.

Mitchell, T. D. and Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 25(6):693–712.

Mönch, E. and Uhlig, H. (2005). Towards a monthly business cycle chronology for the euro area. *Journal of Business Cycle Measurement and Analysis*, 2005(1):43–69.

Moneta, A. and Pallante, G. (2020). Identification of structural var models via independent component analysis: A performance evaluation study. Technical report, LEM Working Paper, Sant'Anna School of Advanced Studies.

Netšunajev, A. and Glass, K. (2017). Uncertainty and employment dynamics in the euro area and the us. *Journal of Macroeconomics*, 51:48–62.

Nordhausen, K., Oja, H., Tyler, D. E., and Virta, J. (2017). Asymptotic and bootstrap tests for the dimension of the non-gaussian subspace. *IEEE Signal Processing Letters*, 24(6):887–891.

Ocasio Cortez, A. (2019). H.Res.109 – 116th Congress (2019–2020): Recognizing the duty of the Federal Government to create a Green New Deal.

OECD (2012). Energy, OECD Green Growth Studies. Technical report, OECD Publishing.

OECD (2018). *Main Economic Indicators, Volume 2018 Issue 12.*

Orea, L., Llorca, M., and Filippini, M. (2015). A new approach to measuring the rebound effect associated to energy efficiency improvements: An application to the US residential energy demand. *Energy Economics*, 49:599–609.

Rausch, S. and Schwerin, H. (2018). Does higher energy efficiency lower economy-wide energy use? *CER-ETH Working Paper*, 18:299.

Saunders, H. D. (2008). Fuel conserving (and using) production functions. *Energy Economics*, 30(5):2184–2235.

Shao, S., Huang, T., and Yang, L. (2014). Using latent variable approach to estimate China's economy-wide energy rebound effect over 1954-2010. *Energy Policy.*

Sorrell, S. and Dimitropoulos, J. (2008). The rebound effect: Microeconomic definitions, limitations and extensions. *Ecological Economics*, 65(3):636–649.

Stern, D. I. (2020). How large is the economy-wide rebound effect? *Energy Policy*, 147:111870.

Stock, J. H. and Watson, M. W. (2016). Dynamic factor models: A brief retrospective. In *Dynamic Factor Models*, pages xv–xx. Emerald Group Publishing Limited.

Turner, K. (2009). Negative rebound and disinvestment effects in response to an improvement in energy efficiency in the UK economy. *Energy Economics*, 31(5):648–666.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419.

Yan, Z., Ouyang, X., and Du, K. (2019). Economy-wide estimates of energy rebound effect : Evidence from China ' s provinces. *Energy Economics*, 83:389–401.

Zhang, J. and Lin Lawell, C.-Y. C. (2017). The macroeconomic rebound effect in China. *Energy Economics*, 67:202–212.

Online Appendix

Appendix A. Data

Energy use: We measure energy use by gross inland consumption (GIC), which covers the amount of energy that is needed to satisfy the total energy demand of a country. Eurostat provides monthly energy data from January 2008 onwards for crude oil (without natural gas liquids), natural gas, and solid fuels[6]. Unfortunately, data on renewable energy sources is only provided as part of the data on total electricity consumption. Including the data on electricity consumption would lead to double counting, as part of total electricity is generated by fossil fuels. Therefore, we decided to only include fossil fuel energy for the European countries. The series are converted from different energy units to Tonne(s) of oil equivalent (toe) and aggregated for each country.[7]

Energy prices: To estimate how energy prices evolved we use indices on harmonized consumer prices (HICP) which measure the changes over time in the prices of consumer goods and services acquired by households. The indices are available for the three different energy carriers (solid, liquid, and gaseous fuels) via Eurostat. To obtain an energy price series we multiply the index with the quarterly end-use energy prices for the industry. To compute the mean cost of the energy carriers in our data we multiply the price series for the different energy carriers with the respective gross inland consumption. Finally, we divide the cost series by the total gross inland consumption.

Figure 2 presents the data series for the three main variables. Note that the data for energy consumption in the U.S. also includes energy from renewables, biomass, and nuclear power generation, which are not included in the European data.

[6]Including data on hard, coke oven and brown coal, peat, oil shale, and oil sands, patent fuels and brown coal briquettes.

[7]For the conversion we used the values from the IEA energy unit converter: https://www.iea.org/classicstats/resources/unitconverter/

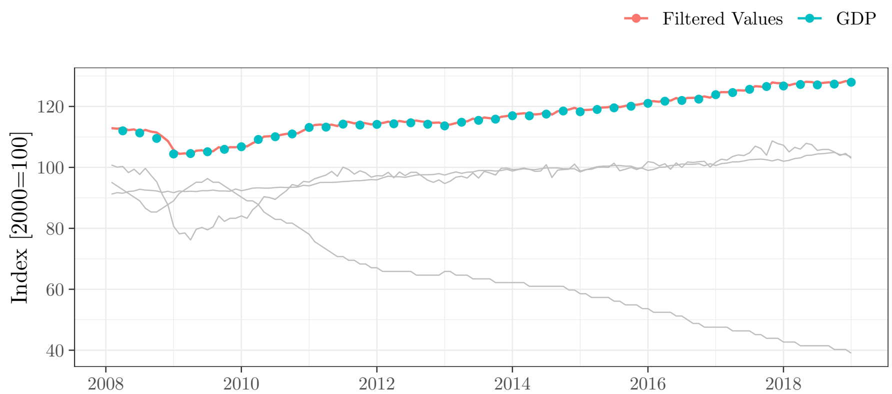

Figure A.7: **Comparison between the interpolated series (red) and the quarterly GDP data (blue) for Germany.** The grey lines in the background depict the evolution of the instrumental series (industrial production, the unemployment rate, and retail trade)

Monthly Economic Indicators: We assume that the factors can be extracted from a large matrix of time series describing the economy. For this purpose, we use the Main Economic Indicator (MEI) database which is developed and maintained by the OECD. This data set presents comparative statistics that provide an overview of recent international economic developments for the European countries we analyzed, covering information on the labor market, national accounts, retail sales, production, construction, prices, finance, international trade, and the balance of payments (OECD, 2018). The latent factors are intended to summarize the main source of variation in the data panel and hence can be interpreted as common driving forces behind different variables of the economy.

Stationarity of the time series in the panel is a precondition for the factor model (Stock and Watson, 2016). To this end, each series of the panel is tested for non-stationarity with the help of an Augmented Dickey Fuller (ADF) test and if the p-value of the test was larger than 0.05, the series is differenced. Then, the time series are tested again and if a time series is still non-stationary, this time series is again differenced. We end up with data sets of the following dimensions for the different countries: Germany (141×56), France (141×53), Italy (141×48), U.K. (141×46) and USA (141×41).

Gross Domestic Production: GDP data is not available on a monthly basis for European countries, therefore we construct monthly series of real GDP based on the encompassing methods proposed by Mönch and Uhlig (2005) and Bernanke et al. (1997). We create a monthly time series for economic activity by combining the available quarterly GDP series and appropriate historical monthly time series. Our approximated GDP series relies on indices capturing employment information, retail trade, and industrial production as instrumental variables.

Although the construction of the monthly GDP series is described in detail in Mönch and Uhlig (2005) and Bernanke et al. (1997), we shall briefly outline the main steps here. We assume that the latent monthly GDP can be explained by correlated high-frequency series using the following dynamic regression framework:

$$(1 - \phi_1 L - \cdots - \phi_p L^p)\, y_t = x_t \beta + u_t \tag{A.1}$$

$$u_t = \rho u_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma^2\right) \tag{A.2}$$

where y_t is the series to be interpolated, i.e. unobserved monthly GDP, and x_t are exogenous covariates that exhibit a high correlation with y_t. The monthly GDP series as well as the regression residuals u_t are assumed to follow an AR-process of lag order p (We use $p = 1$ for simplicity).

Furthermore, the mean of three consecutive monthly GDP values shall equal exactly one third of the observed quarterly GDP value and y_t equals zero for the months that information on GDP is missing. Hence, y_t and y_t^+ are connected by the following measurement equation:

$$y_t^+ = \frac{1}{3} \sum_{i=0}^{2} y_{t-i}, \quad t = 3, 6, 9, \ldots \tag{A.3}$$

$$y_t^+ = 0 \quad \text{otherwise} \tag{A.4}$$

The relationships between the observable and the latent series can be encompassed in the following state-space form:

$$y^{+} = \boldsymbol{H}_t' \xi_t \tag{A.5}$$

$$\xi_t = \begin{pmatrix} \phi & 0 & 0 & \rho \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \rho \end{pmatrix} \begin{pmatrix} y_{t-1} \\ y_{t-2} \\ y_{t-3} \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} x_t'\beta \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \varepsilon_t \\ 0 \\ 0 \\ \varepsilon_t \end{pmatrix} \tag{A.6}$$

where the matrix H_t varies as specified:

$$\mathrm{H}_t' = \begin{cases} \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 \end{bmatrix}, & t = 3, 6, 9, 12, \ldots, T \\ \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}, & \text{otherwise} \end{cases} \tag{A.7}$$

According to Issler and Notini (2016) this form of interpolation offers two considerable advantages: first, the form is especially appealing because it is able to incorporate different model specifications used for interpolation in a unified framework. More precisely, depending on the restrictions on the parameters ϕ and ρ several models can be estimated amongst them the specifications of Chow and Lin (1971), Fernandez (1981), and Mitchell and Jones (2005). Second, we ensure that the aggregate of each three-months of interpolated GDP data is equal to measured quarterly GDP.

Following Mönch and Uhlig (2005) we treat the monthly GDP values as latent states which are estimated using the Kalman filter. We estimate the parameters ϕ, ρ and the variance of u_t via Maximum Likelihood.

The choice of the related series, x_t, is crucial for the results of the interpolation procedure as they provide the signals extracted for the monthly estimates. The series do not only have to be available on a monthly frequency but also be highly correlated with the quarterly GDP series. Monthly data is in general quite scarce for European countries. Natural candidates for this kind of business cycle analysis, and also the series used by Mönch and Uhlig (2005) and Issler and Notini (2016) are industrial production, retail sales, income, exports, and employment. We decided on the combination of instrumental variables conditional on data availability as well as on the measures of fit in first differences for the filtered and smoothed GDP estimates (see Table A.3). The latter is

Country	Related Series	$R^2_{filtered}$	$R^2_{smoothed}$
Italy	IP, Emp, Ret	0.29	0.65
UK	IP, Emp, Ret	0.62	0.88
France	IP, Emp, Ret	0.73	0.92
Germany	IP, Emp, Ret	0.59	0.87

Table A.3: Related series used for interpolation (IP=Industrial Production, Emp=Employment, Ret=Retail sales, CPI=Consumer Price Index) and the R^2 goodness of fit statistics as measures of interpolation quality

calculated using the following R^2:

$$R^2_{diffs} = \frac{\mathrm{Var}\left(\Delta y_{iT}\right)}{\mathrm{Var}\left(\Delta y_{iT}\right) + \mathrm{Var}\left(\Delta u_{iT}\right)} \tag{A.8}$$

The results in Table A.3 as well as the visual inspection of the plots in A.8 indicate that the interpolation is reasonably accurate. The smoothed estimates of the interpolated GDP serve as our GDP sequence.

Additionally, we estimated a GDP series that is approximated by IP only and used it as a robustness check in the analysis. The results did not differ significantly.

Figure A.8: **Comparison between the interpolated series (red) and the quarterly GDP data (blue) for European countries.** The grey lines in the background depict the evolution of the instrumental series industrial production, unemployment, and retail trade indices.

Appendix B. Forecast Error Variance Decomposition

The FEVD are a measure of the impact of an impulse on the modeled variables, and, specifically, on their predicted values, since it studies how much of the variance of the error one makes in forecasting a variable x_t (in other words, the prediction mean squared error of x_t) at different time horizons, is accounted for by the different impulse shocks. If an impulse accounts for most of the forecast error variance of a specific variable x_t at most of the time horizons, this is a good evidence that the impulse shock should be labeled as shock for x_t.

While we focus on identifying the energy efficicency impulse (first column of Figure B.9), column two and three also show the FEVD for what we labelled the energy price impulse and the GDP impulse. This holds even if in some countries, e.g. UK, US and France, energy use receives some impact from what we label the energy price impulse, whose percent of its prediction mean squared error range from about 10% to 20%. For most of the countries, it is obvious that labeling the impulse which has the greatest impact on price (on GDP) as the price (GDP) impulse is also correct. However, this is not obvious for Italy, where two impulses impact almost equally on price and GDP. As mentioned, the identification of these two shocks is not relevant for the estimate of the rebound effect.

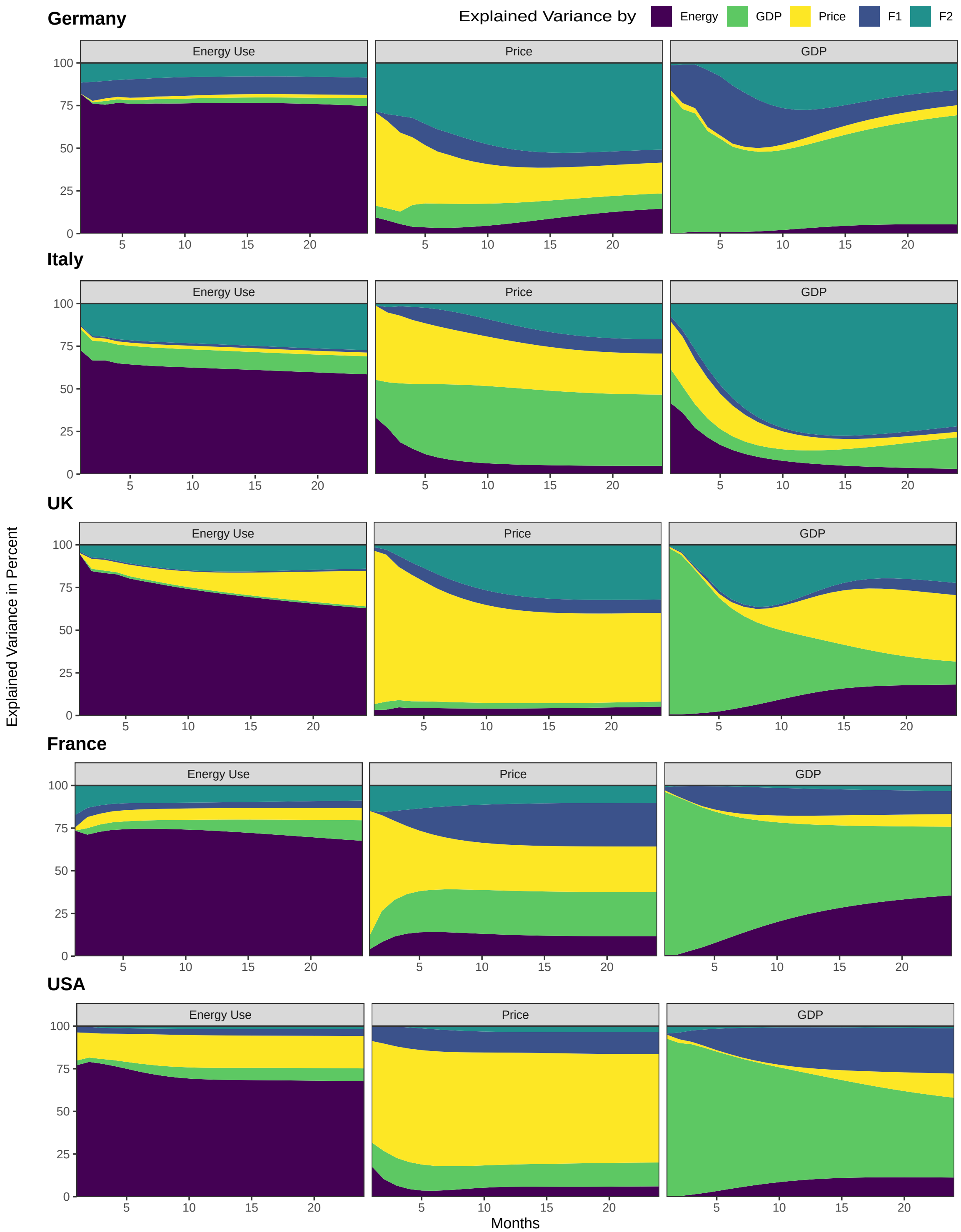

Figure B.9: **Forecast error variance decomposition.** The decomposition, for each variable and country, shows the percent (y axis) of h-months (x axis) ahead forecast error variance (prediction mean squared error) explained by five impulse shocks, which we label as the Energy, GDP, Price, Factor 1, Factor 2 impulses.

Appendix C. Economic plausibility of the estimated S-FAVAR models

We analyse the impulse response functions for the three main variables. We would expect the energy efficiency shock to have a negative contemporaneous effect on energy and a positive or zero effect on GDP in the long run (because TFP increases and consumers have more real income after an efficiency improvement), and a negative effect on the price of energy. Looking at Figure C.10, all countries have the negative effects on energy use, which reduce over time. The UK, Italy, and Germany have a positive impact on GDP and the US the most negative though the confidence interval includes zero. France has a negative mean but a rising trend.[8]

We would expect the GDP shock to have a large positive effect on GDP, which we see in all countries. Standard demand theory and cross-country studies suggest that a GDP shock should increase energy use. However, structural change associated with economic growth could lower energy use. This is what we see in all countries apart from the US. If we assume a standard supply and demand setting and assume that growth only moves the demand curve, then if energy use is reduced by the GDP shock it should reduce the price of energy and vice versa. There is a mixed picture with the GDP shock raising the price of energy in the UK (and insignificantly in Italy) despite reducing energy use and increasing the price of energy in the US and lowering it in France and Germany as we would expect.

We notice that the GDP shock appears to be permanent in most countries, which is what we would expect for an increase in TFP or population.

By contrast, price shocks have temporary positive effects on the price of energy. The price shock mostly has a negative or zero effect on energy use, though in the US the initial effect is large and positive. The expected negative or zero effects on GDP are present except in the UK though the latter is not statistically significant. Speculatively a positive effect could happen in an oil producing country. But we do not see this in the US.

In summary, we mostly see the expected theoretical effects though there are notable exceptions.

[8]For the impulse response functions for the individual countries, refer to Figure C.11-C.15

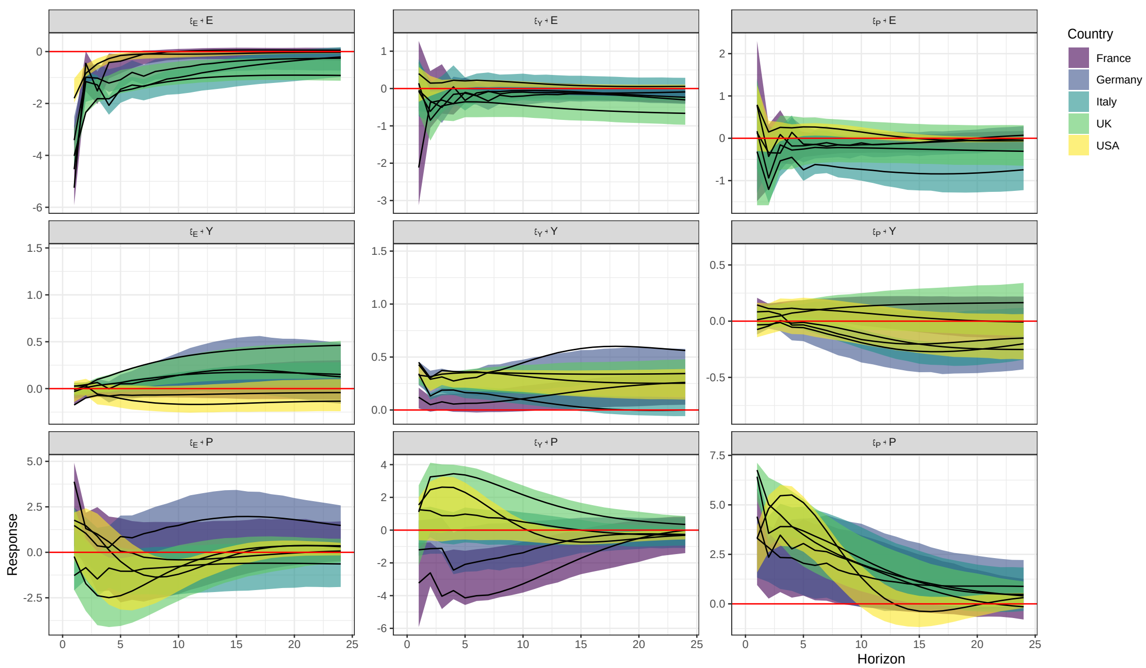

Figure C.10: **Impulse Response Functions based on the distance covariance approach.** The first column shows the effect of the energy efficiency impulse (ε_E) on energy use (E), GDP (Y) and energy price (P). Analogously, the second and third columns show the effects of the GDP (ε_Y) and energy price (ε_P) impulses. The shaded areas indicate 68% confidence intervals.

Table C.4: Contemporaneous reaction of variables to different impulses

	ϵ_E	ϵ_Y	ϵ_P	ϵ_{F_1}	ϵ_{F_2}
Germany					
e_t	-3.41	-0.05	0.13	-0.94	-1.27
	(-3.7, -1.32)	(-1.04, 1.19)	(-1.82, 1.74)	(-1.53, 0.96)	(-3.12, 0.56)
y_t	0.03	0.43	0.08	-0.17	-0.06
	(-0.16, 0.14)	(0.14, 0.47)	(-0.18, 0.37)	(-0.34, 0.09)	(-0.3, 0.17)
p_t	1.47	-1.2	3.32	-0.34	2.46
	(-0.92, 3.4)	(-2.62, 1.08)	(0.83, 4.43)	(-1.98, 1.23)	(-0.67, 4.26)
France					
e_t	-5.25	-2.12	0.78	0.55	-2.18
	(-6.05, -1.85)	(-5.15, 1.91)	(-2.43, 4.11)	(-1.53, 2.23)	(-4.23, 1.29)
y_t	-0.17	0.12	0.14	0.04	0.07
	(-0.23, -0.02)	(-0.03, 0.24)	(-0.01, 0.23)	(-0.09, 0.12)	(-0.05, 0.19)
p_t	3.88	-3.23	4.4	0.11	0.66
	(-0.48, 5.88)	(-6.43, 0.33)	(-0.12, 6.34)	(-2.46, 2.85)	(-2.21, 2.96)
Italy					
e_t	-4.53	-0.07	-0.32	0.97	-0.41
	(-4.61, -3.51)	(-1.37, 1.48)	(-1.47, 2.06)	(-1.6, 1.87)	(-1.52, 1.33)
y_t	-0.03	0.34	-0.03	0.03	0.01
	(-0.13, 0.08)	(0.27, 0.34)	(-0.09, 0.13)	(-0.08, 0.13)	(-0.08, 0.09)
p_t	-1.28	1.25	6.42	0.76	-1.04
	(-3.08, 2.04)	(-1.94, 2.27)	(4.96, 6.71)	(-2.13, 2.37)	(-2.16, 1.61)
United Kingdom					
e_t	-4.02	0.14	0.17	-1.01	1.91
	(-4.51, -1.46)	(-1.3, 1.21)	(-2.44, 2.2)	(-1.58, 1.22)	(-1.19, 3.8)
y_t	0.03	0.33	0.01	-0.03	0.03
	(-0.11, 0.11)	(0.21, 0.32)	(-0.14, 0.17)	(-0.2, 0.08)	(-0.13, 0.13)
p_t	-0.23	1.11	6.76	-0.1	-3.19
	(-3.43, 3.58)	(-3.22, 3.63)	(3.41, 7.36)	(-2.9, 2.06)	(-5.46, 1.57)
United States					
e_t	-1.8	0.4	0.79	-0.38	-0.09
	(-2.01, -0.68)	(-0.66, 1.04)	(-0.56, 1.76)	(-0.86, 0.6)	(-0.8, 0.64)
y_t	-0.01	0.45	-0.07	0.04	-0.08
	(-0.22, 0.17)	(0.22, 0.45)	(-0.32, 0.22)	(-0.15, 0.19)	(-0.25, 0.17)
p_t	1.76	1.54	3.31	1.27	0.07
	(-0.91, 3.43)	(-1.09, 3.09)	(1.12, 4.1)	(-0.96, 3.09)	(-2.53, 1.58)

Notes: Bootstrapped 95% confidence intervals in parentheses.

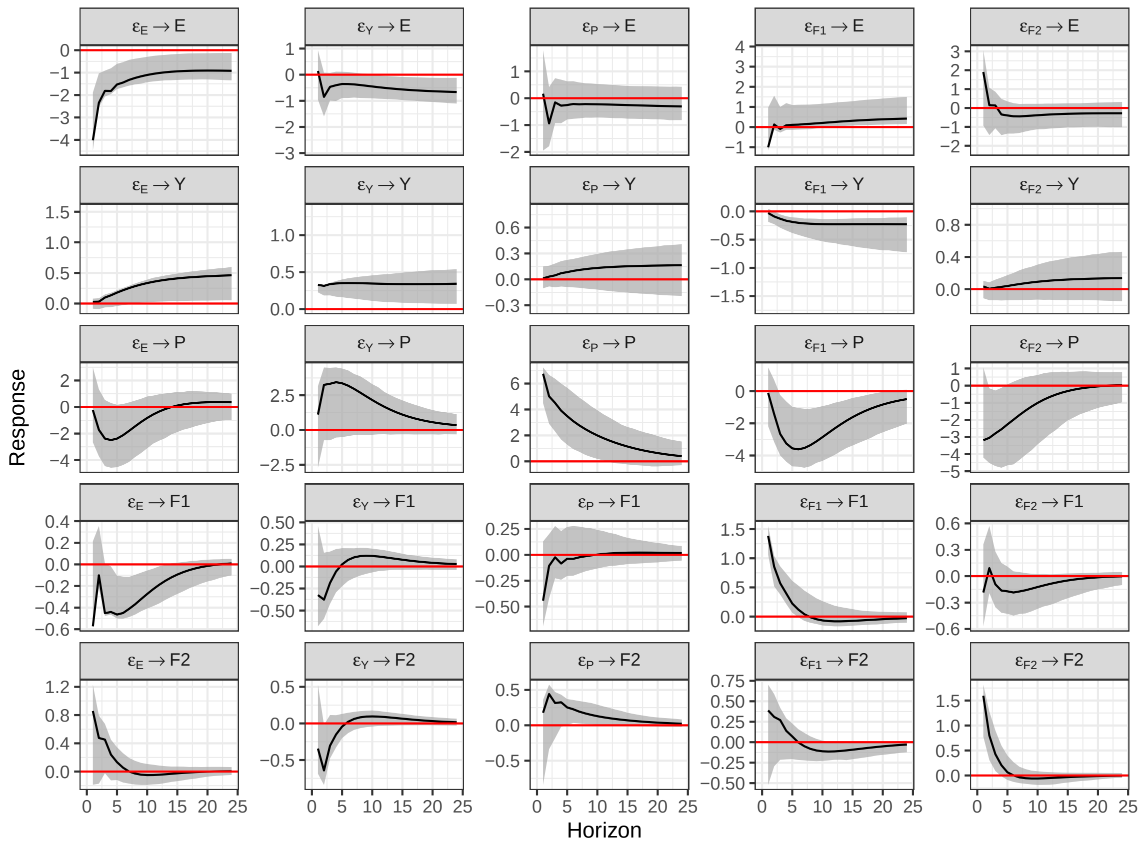

Figure C.11: **Impulse Response Functions for the U.K.**, identified with the distance covariance approach. The grey-shaded areas indicate 68% confidence intervals

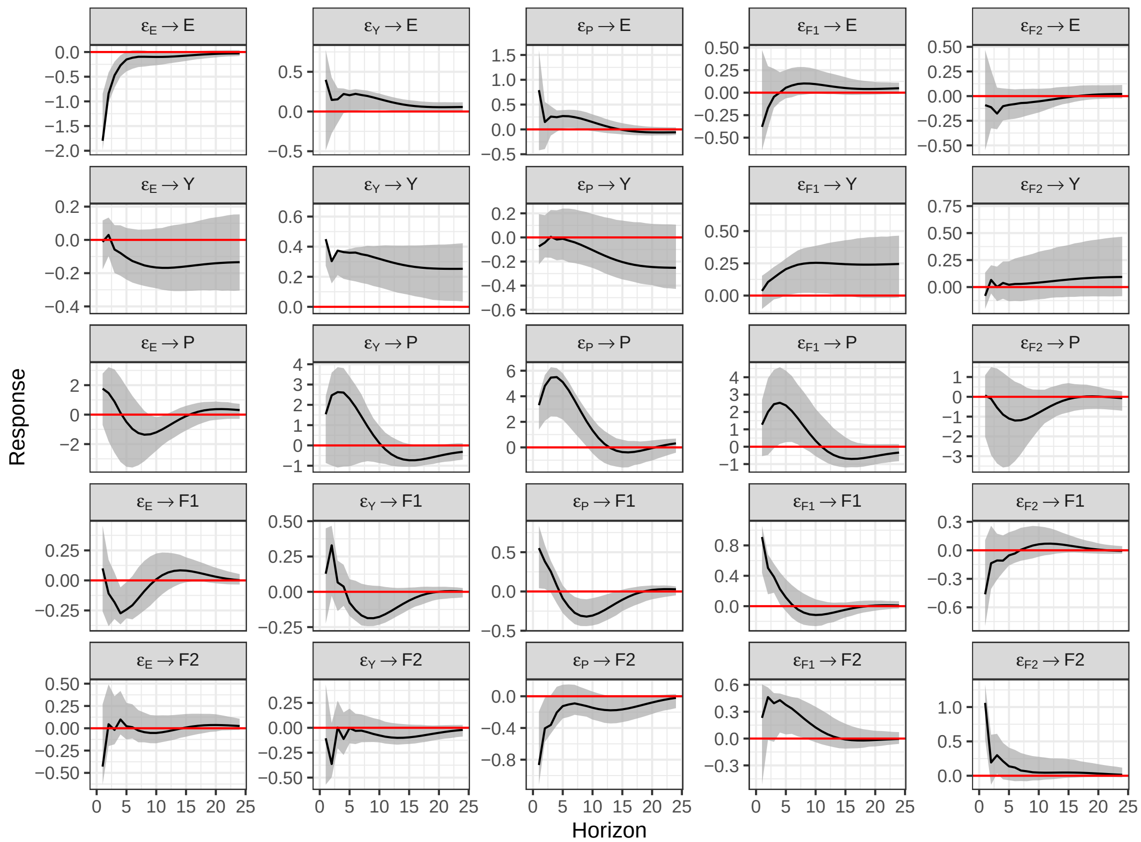

Figure C.12: **Impulse Response Functions for the US**, identified with the distance covariance approach. The grey-shaded areas indicate 68% confidence intervals

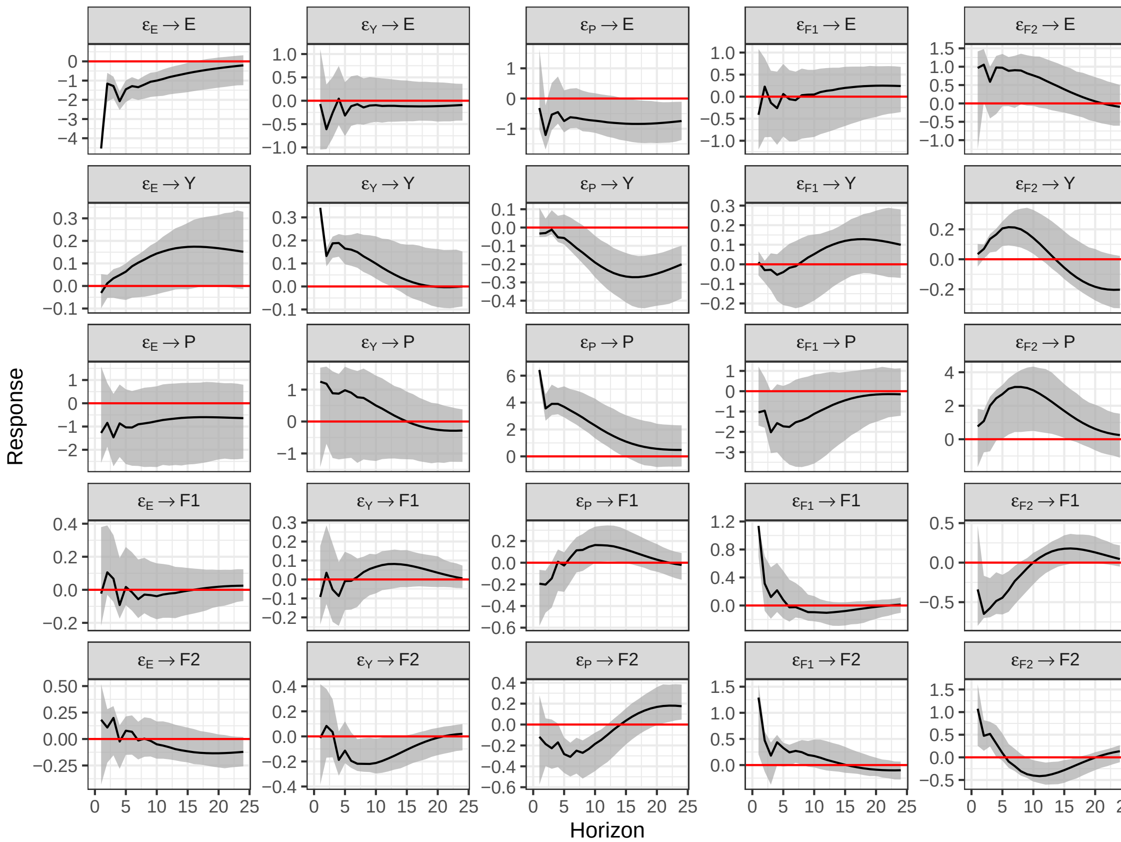

Figure C.13: **Impulse Response Functions for Italy**, identified with the distance covariance approach. The grey-shaded areas indicate 68% confidence intervals

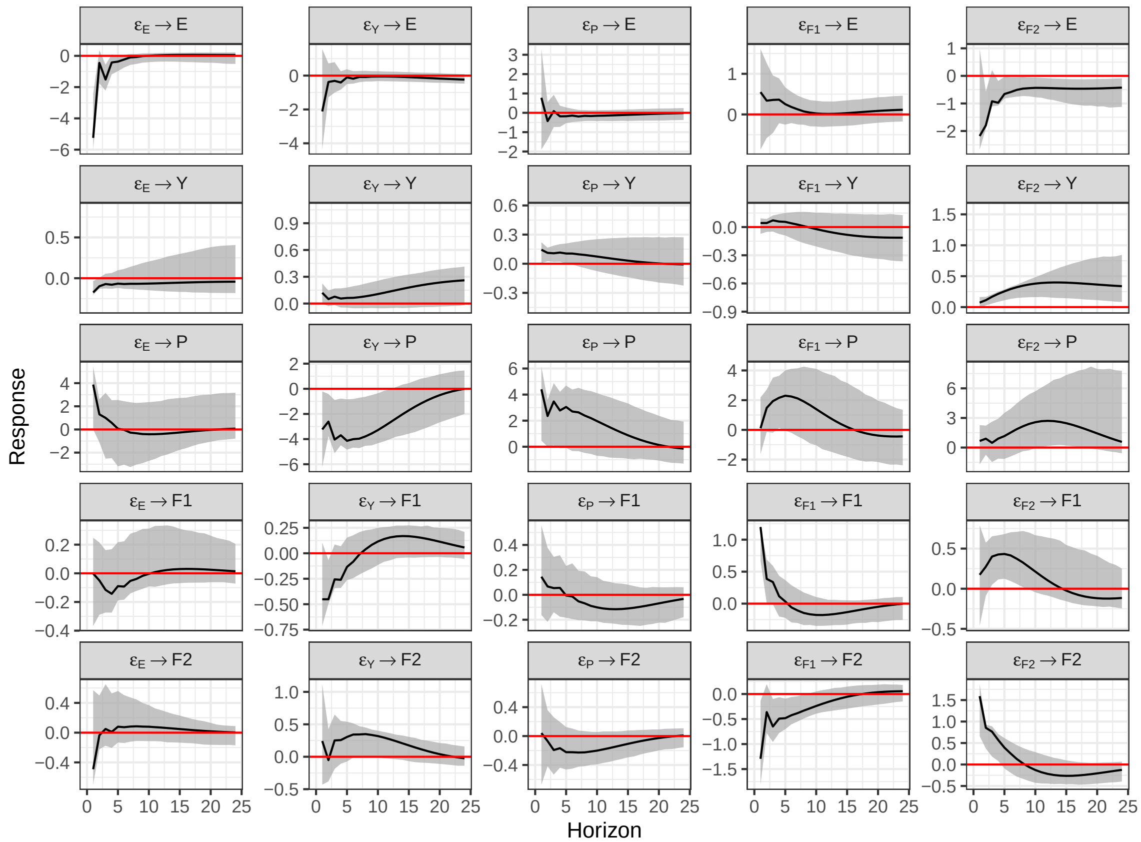

Figure C.14: **Impulse Response Functions for France**, identified with the distance covariance approach. The grey-shaded areas indicate 68% confidence intervals

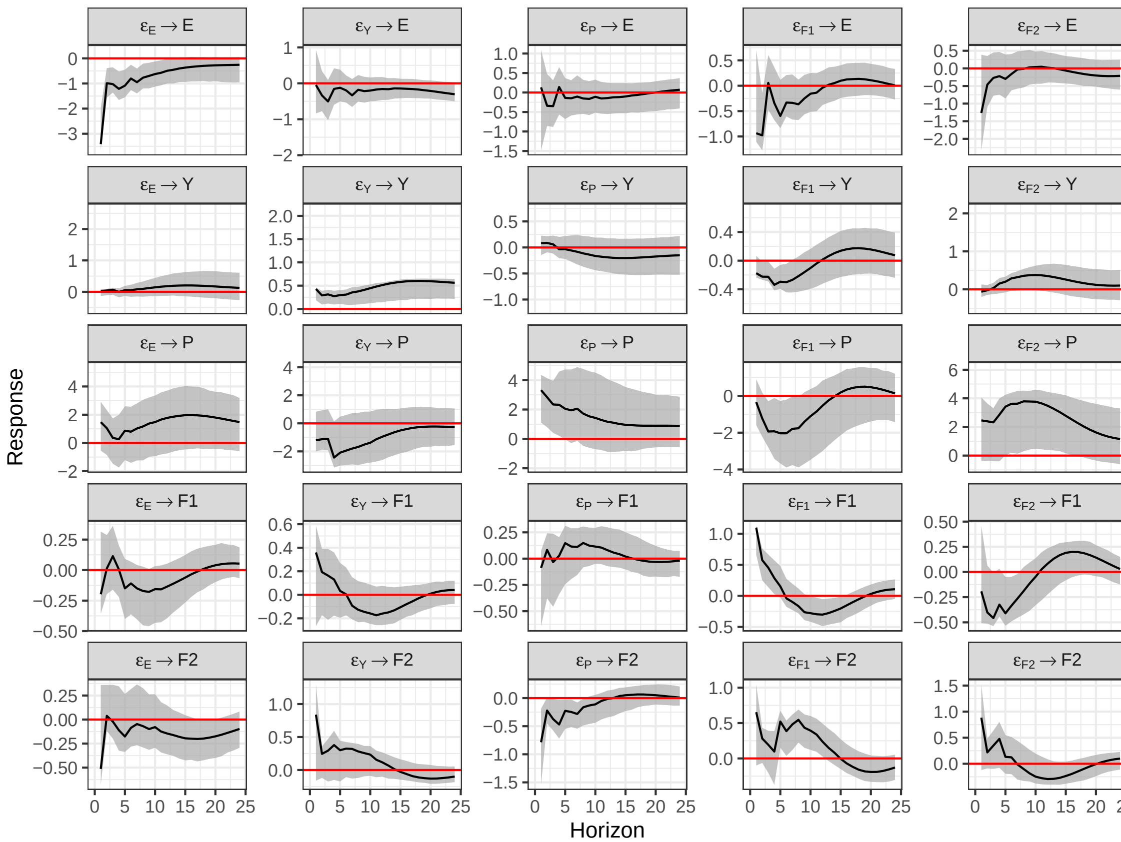

Figure C.15: **Impulse Response Functions for Germany**, identified with the distance covariance approach. The grey-shaded areas indicate 68% confidence intervals

Appendix D. Additional Figures and Tables

Table D.5: Lag length selection based on the AIC. Comparison of a maximal lag length of 6 and 12.

	country	lag.max = 6	lag.max = 12
1	Germany	4	4
2	France	2	2
3	Italy	3	3
4	UK	2	2
5	USA	2	2

Table D.6: Estimated rebound effects for different identification methods

Country	Months	Distance covariance	Non-gaussian ML	Choleski
Germany	3	0.70	0.72	0.72
	6	0.77	0.76	0.77
	12	0.86	0.87	0.88
	24	0.93	0.92	0.91
France	3	0.71	0.69	0.71
	6	0.96	0.93	0.90
	12	1.01	0.99	0.96
	24	1.01	1.00	0.94
Italy	3	0.72	0.73	0.69
	6	0.72	0.73	0.68
	12	0.82	0.86	0.79
	24	0.95	1.02	0.95
UK	3	0.55	0.63	0.64
	6	0.64	0.74	0.77
	12	0.75	0.84	0.85
	24	0.77	0.87	0.87
USA	3	0.74	0.73	0.73
	6	0.94	0.93	0.90
	12	0.95	0.95	0.95
	24	0.98	0.99	1.00

Notes: Comparison of estimated rebound effects, using the ICA approaches (distance covariance and non-Gaussian Maximum Likelihood) and a classic Choleski decomposition with the causal order: $y \rightarrow e \rightarrow p \rightarrow F_1 \rightarrow F_2$.

Figure D.16: **The 15 highest factor loadings considering the first two factors.** Notes: "ExRates" stands for exchange rates, "IntRates" for interest rates, "LI" for leading indicator, "PPI" for producer price index and "HCPI" for harmonized consumer price index.

We statistically evaluate the number of Gaussian components based on a fourth order blind identification (FOBI) to the reduced-form model disturbances $\hat{u}_t$ (Nordhausen et al., 2017). This procedure evaluates the residuals under the null hypothesis of $K-k$ non-Gaussian components, where K is the total number of components and k is the number of Gaussian components. Except for Germany, this hypothesis can be rejected for all countries. For all countries, we can not reject the null hypothesis, which indicates that all components except for one are non-Gaussian ($k=4$). For France and the UK the hypothesis of $k=3$ cannot be rejected, hinting at two Gaussian components and for the US and Italy the hypothesis of $k=2$ and therefore of three Gaussian components. Last, for Germany all components could be considered Gaussian according to the test results of the IC test. The fact that there is more than one Gaussian component in the model is also reflected by the component-wise tests for the time series (see Table D.7).

Table D.7: Component-wise normality tests for the different time series

country	Variable	SW	SF	JB	IC Test, k = 1	k = 2	k = 3
Germany	E	0.01	0.01	0.00	0.562	0.631	0.9165
	Y	0.69	0.56	0.91	-		
	P	0.26	0.32	0.50	-		
	F1	0.06	0.02	0.00	-		
	F2	0.33	0.27	0.49	-		
France	E	0.73	0.47	0.28	0.003	0.042	0.5642
	Y	0.33	0.24	0.19	-		
	P	0.80	0.77	0.93	-		
	F1	0.00	0.00	0.00	-		
	F2	0.10	0.04	0.01	-		
Italy	E	0.34	0.13	0.09	0.023	0.41	0.5867
	Y	0.61	0.51	0.59			
	P	0.15	0.10	0.13			
	F1	0.00	0.00	0.00			
	F2	0.27	0.17	0.46			
UK	E	0.00	0.00	0.00	0.001	0.023	0.9745
	Y	0.00	0.00	0.00			
	P	0.02	0.01	0.00			
	F1	0.24	0.16	0.30			
	F2	0.09	0.06	0.26			
USA	E	0.01	0.01	0.00	0.025	0.219	0.6557
	Y	0.16	0.07	0.07			
	P	0.24	0.40	0.32			
	F1	0.00	0.00	0.00			
	F2	0.21	0.10	0.10			

Notes: SW = Shapiro-Wilk, SF = Shapiro-Francia, JB = Jarque-Bera and IC stands for an independence based boostrappping test, based on fourth-order blind identification tests from the package `ICtest` (Nordhausen et al., 2017).

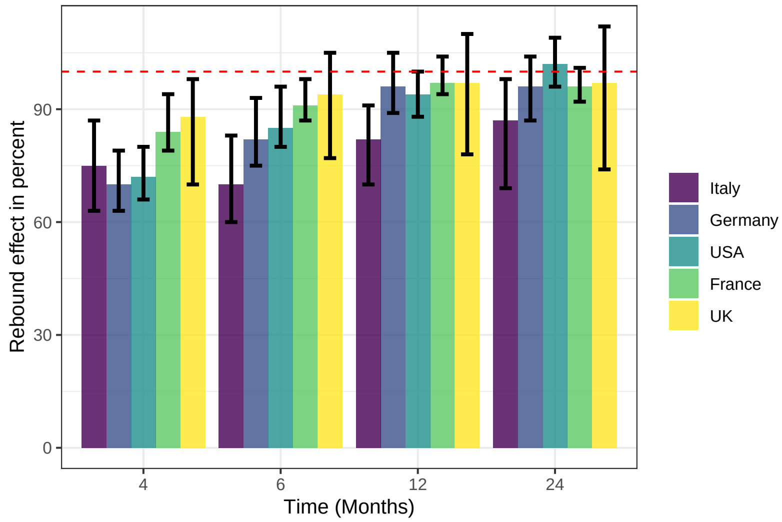

Figure D.17: **Estimated rebound effects for a model including 3 factors**. Error bars represent 90% confidence intervals.

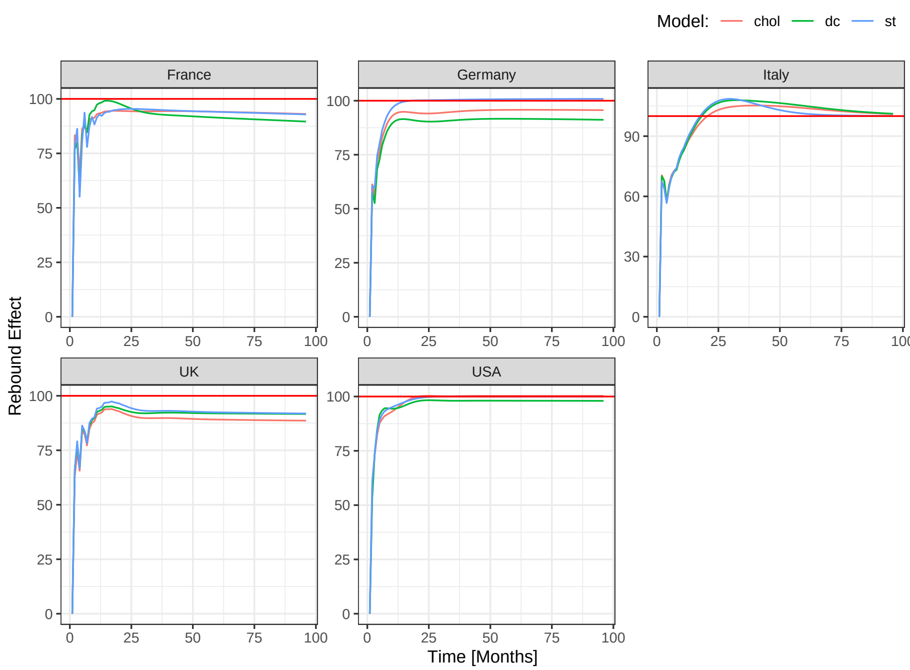

Figure D.18: **Impulse-response functions resulting from various identification strategies.** The identification via smooth transitions of covariances (st), the dictance-covariance approach building on ICA and the Choleski decomposition (chol) all show comparable results for the depicted impulse-response to an energy efficiency shock