Estimating and Testing the Multicountry Endogenous Growth Model[*]

Stef De Visscher[1,2], Markus Eberhardt[3,4], and Gerdie Everaert[1]

[1]*Ghent University, Belgium*
[2]*Research Foundation Flanders (FWO), Belgium*
[3]*University of Nottingham, U.K.*
[4]*Centre for Economic Policy Research, U.K.*

March 24, 2020

Abstract: We estimate Cobb-Douglas production functions that parameterize unobserved total factor productivity as a global technology process interacted with country-specific absorptive capacities. In contrast to the existing literature we do not require proxies for these absorptive capacities but instead estimate them as time-varying stochastic processes. Our implementation allows us to test the contrasting predictions of alternative growth models and our results for a panel of advanced economies support the multicountry endogenous growth model in that an enhancement in absorptive capacity raises a country's long-run productivity level but not its growth rate. This finding is confirmed in an extended model where we allow a set of policy variables (financial development, human capital, competition policy, and knowledge stock) to affect absorptive capacity, none of which induce permanent growth effects. The proxies for financial development and knowledge stock stand out for their significant level effects.

Highlights:

- TFP is a global technology process with country-specific absorptive capacities;

- Baseline specification: absorptive capacities are time-varying stochastic processes;

- Extension: absorptive capacities are also a function of observed variables;

- Absorptive capacity investments raise the level of TFP but no its growth rate.

---

[*]Correspondence: Markus Eberhardt, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, United Kingdom. Email: markus.eberhardt@nottingham.ac.uk

1 Introduction

Output per capita shows enormous and persistent differences across countries. As variations in factor inputs are unable to explain these differences, there is an important role for disparities in total factor productivity (TFP). The relative importance of TFP vis-à-vis factor accumulation for economic growth has occupied economists not least since Tinbergen (1942), Abramovitz (1956) and Solow (1956). The traditional neoclassical growth model leaves TFP unexplained and, in the famous words of Edward Denison (1967, 282), assumes that "[b]ecause knowledge is an international commodity, [we] should expect the contribution of advances of knowledge... to be about the same size in all the countries." Much of the early endogenous growth literature relates a country's TFP growth rate to its research efforts. The implication of this type of single-country model without international spillovers is long-run divergence of incomes across countries, given that countries with higher research efforts should permanently leave all others behind. However, this prediction seems at odds with the available data on cross-country income differences. Situated in-between the neoclassical and single-country endogenous growth models are the multicountry endogenous growth models of Eaton and Kortum (1999), Howitt (2000), and Klenow and Rodríguez-Clare (2005), among others. These models are in the neoclassical tradition, in that economic growth in steady state is equal for all countries (thanks to knowledge spillovers), but they also capture some of the defining features of endogenous growth models since differences in innovation efforts among countries are important for steady state *levels* of TFP.[1] In this strand of the literature, TFP is given the interpretation of successfully assimilated ('absorbed') global technology. What unites concepts such as absorptive capacity and alternatives — e.g. social capability (Abramovitz, 1986) — is the notion that despite the designation of knowledge as a public good or being in the public domain, technological catch-up is by no means guaranteed, but requires considerable efforts and investments (see Parente and Prescott, 1994, 2002; Aghion and Jaravel, 2015, among others).[2] Related work by Benhabib and Spiegel (1994, 2005) builds on and extends the Nelson and Phelps (1966) model which assigns a central role for human capital in the knowledge diffusion process: instead of playing the role of a (mere) factor input such as in Mankiw et al. (1992), among many others, human capital is suggested to facilitate knowledge diffusion. Here the implication is that an increase in average education levels (proxying absorptive capacity) leads to a permanent increase in TFP *growth*.[3]

In the empirical growth literature there is a long tradition of quantifying 'foreign' elements of TFP by assuming specific channels through which international knowledge spillovers can occur and/or pinpoint-

---

[1]We can further extend the notion expressed in Howitt (2000) that countries which develop their absorptive capacity (he uses R&D) converge, whereas those where incentives to develop these characteristics are missing will fall behind.

[2]For a detailed discussion of the origins of absorptive capacity see Fagerberg et al. (2010). In this article we use knowledge spillovers synonymously with 'technology spillovers' or more broadly the assimilation of ideas and innovations developed in other countries. Technology is used interchangeably with productivity, knowledge and TFP.

[3]For the diffusion function of human capital Benhabib and Spiegel (2005) provide two alternative specifications, although in our advanced country context it would seem unlikely that 'follower' human capital stock could be too low and thus leading to TFP divergence as postulated in their 'logistic model.'

ing country characteristics deemed synonymous with absorptive capacity. The most prominent channels are arguably the patterns of international trade, foreign direct investment and international migration (Coe and Helpmann, 1995; Pottelsberghe and Lichtenberg, 2001; Madsen, 2007; Acharya and Keller, 2009; Bahar et al., 2014; Bosetti et al., 2015; Fracasso and Marzetti, 2015, see also Keller (2004, 2010) for detailed surveys). Human capital (Griffith et al., 2004; Benhabib and Spiegel, 2005; Madsen et al., 2010; Ertur and Musolesi, 2017) and investment in R&D (Aghion and Howitt, 1998; Griffith et al., 2003; Mancusi, 2008; Lu et al., 2017) are frequently employed as proxies for absorptive capacity. While *a priori* all of these factors and channels are likely to be relevant to capture the discovery and assimilation of ideas developed elsewhere, estimates are biased if the proxies included are correlated with other, omitted determinants. This criticism points to recent efforts to quantify the quality of management or productivity-enhancing investment in *intangible* capital, of which formal R&D is only one element (see e.g. Acharya, 2016; Bloom et al., 2016; Corrado et al., 2017). Moreover, Eberhardt et al. (2013) argue that the empirical literature on knowledge spillovers following the seminal contribution by Coe and Helpmann (1995) suffers from an omitted variable bias induced by spillovers and common shocks with heterogeneous impact. They show that private returns to R&D are dramatically lower once this type of cross-sectional correlation is taken into account.

The present paper takes a different approach. We specify a Cobb-Douglas production function that parameterizes unobserved total factor productivity as a global technology process interacted with country-specific absorptive capacity that varies stochastically over time. We demonstrate that our econometric specification is a generalization of the multicountry endogenous growth model referred to above. To identify unobserved global technology, we extend Pesaran's (2006) common correlated effects (CCE) approach to a setup where factor loadings evolve according to random walks. Written in state space representation, the model can be estimated using Bayesian Markov Chain Monte Carlo (MCMC) sampling methods. This produces time-varying country-specific indices for absorptive capacity that do not rely on *ad hoc* proxies or impose explicit channels through which knowledge spills over but encompass all of these. We show how the relevance of time variation in absorptive capacities can be tested as part of the implementation using the stochastic model specification search of Frühwirth-Schnatter and Wagner (2010). This allows us to test the differential implications of alternative theoretical growth models about whether improvements in absorptive capacity have growth or merely level effects.

We also estimate an extended model including observed country characteristics that are believed to affect absorptive capacity. This more general specification maintains the flexibility of the baseline approach yet allows absorptive capacity to be driven by both an unobserved stochastic component and observed (policy) variables. We will focus on a set of variables highlighted in the recent Schumpeterian growth literature dominating the current debate on policy for economic growth: financial development, human capital, competition policy and the R&D knowledge stock.

We estimate our baseline model for a panel of 31 countries covering the period 1953-2014. The choice of sample countries is driven by our desire to capture the technology diffusion and absorption process *in today's advanced economies*, including economies such as the United States which have been at the frontier throughout this period, as well as others such as South Korea or Taiwan, which have experienced impressive catch-up growth. Our first finding is that there is relevant time variation in absorptive capacities. Especially countries like Ireland, South Korea and Taiwan have been able to increase their ability to assimilate foreign knowledge over the sample period. Second, changes in absorptive capacity are *not* found to have permanent growth effects, but merely to affect the level of TFP.

Our third finding relates to the country-specific evolution of absorptive capacity, which can be squared with policy levers to promote innovation and technical change:[4] We estimate an extended model using a reduced panel of 21 countries for which we have indicators of financial development, human capital, competition policy/regulation, and R&D expenditure over the period 1970-2009. In line with the baseline model, none of these variables induces permanent *growth* effects. In our results financial development has a strong positive effect on absorptive capacity, yet is subject to diminishing returns, echoing the "too much finance" argument in recent work on economic growth (Arcand et al., 2015). Moreover, R&D stock is found to contribute substantially to our measure of absorptive capacity in line with arguments for the 'two faces of R&D' (Griffith et al., 2004).

The remainder of this study is structured as follows. Section 2 sets out the baseline empirical specification, demonstrates how it can be squared with a multicountry endogenous growth model and outlines the Bayesian simulation-based method that will be used to estimate the model. The data along with the empirical results are discussed in Section 3. In Section 4 we present and estimate an extended model that allows absorptive capacity to be driven by both an unobserved stochastic component and observed variables. Section 5 concludes.

---

[4]In **??** we cherry-pick three economies (Ireland, Sweden, and Japan) with very different trajectories and provide a coherent narrative in this vein.

2 Empirical specification and implementation

We present a factor-augmented Cobb-Douglas production function with time-varying absorptive capacity and suggest a CCE approach to identify unobserved global technology. Transformed into a state space model our empirical model can be estimated using MCMC simulation methods, which can further be employed to test whether changes in absorptive capacity induce level or growth effects.

2.1 Empirical model

We model output in country $i = 1, \ldots, N$ at time $t = 1, \ldots, T$ using a Cobb-Douglas production function with constant returns to scale

$$Y_{it} = \Lambda_{it} K_{it}^{\beta_i} L_{it}^{1-\beta_i} e^{\epsilon_{it}}, \qquad \text{with} \quad 0 < \beta_i < 1 \quad \forall i, \tag{1}$$

where Y_{it} is real GDP, K_{it} is the real aggregate (private and public) capital stock, L_{it} is total hours worked and ϵ_{it} is a zero-mean stationary error term uncorrelated across countries. The country-specific elasticity of output with respect to capital β_i is approximated by country i's share of GDP that is not earned by labor (i.e., one minus the economy-wide labor income share). Equation (1) implies that unobserved TFP, Λ_{it}, is defined to capture the stock of intangible technology and knowledge, and the effects of human capital, among other factors.

Common factor structure of time-varying absorptive capacity

Building on an established strand of the literature that considers a country's TFP to be the successful assimilation of global technology (Parente and Prescott, 1994, 2002; Alfaro et al., 2008), we parameterize Λ_{it} using a common factor framework

$$\Lambda_{it} = \mathcal{A}_{it} F_t^{\vartheta_{it}}, \tag{2}$$

where F_t is a common factor that we interpret as representing the worldwide available technology and knowledge stock while $\mathcal{A}_{it}$ and ϑ_{it} capture country-specific endowments, institutions, investments and policies that determine how much of F_t is successfully assimilated (henceforth 'absorptive capacity'). Substituting equation (2) in (1), dividing by hours worked L_{it} and taking logarithms yields the following expression for log TFP, $\lambda_{it} = ln\left(\Lambda_{it}\right)$

$$(y_{it} - \beta_i k_{it}) = \lambda_{it} = a_{it} + \vartheta_{it} f_t + \epsilon_{it}, \tag{3}$$

where $y_{it} = ln\left(Y_{it}/L_{it}\right)$, $k_{it} = ln\left(K_{it}/L_{it}\right)$, $a_{it} = ln\left(\mathcal{A}_{it}\right)$ and $f_t = ln\left(F_t\right)$.

Impact of changes in absorptive capacity: shifts in TFP levels versus growth rates

The empirical specification in equation (3) is closely related to that of Eberhardt et al. (2013), who use a common factor framework with time-invariant parameters, and Everaert et al. (2014), who allow absorptive capacity to vary as a function of fiscal policy variables. Our main methodological contribution is to allow for a flexible evolution in a_{it} and ϑ_{it}, and hence in absorptive capacity, that is able to capture all relevant channels through which knowledge spills over. This empirical specification for Λ_{it} is a generalization of the multicountry endogenous growth model outlined in Eaton and Kortum (1999), Howitt (2000) and Klenow and Rodríguez-Clare (2005), among others, where differences in policies or other country efforts to improve absorptive capacity generate shifts in TFP levels rather than growth rates. The main result of this type of model is that in the long-run all countries share a common growth rate equal to the growth rate of global TFP — a result they empirically motivate by demonstrating that countries with high investment rates typically have higher levels of wealth rather than higher growth rates. In our empirical specification this equates to setting $\vartheta_{it} = 1$. In order to allow for an endogenous type of growth where country-specific characteristics, research efforts or policies can have permanent growth effects (Benhabib and Spiegel, 2005) we allow for the possibility that $\vartheta_{it} \neq 1$ and that it varies across countries and time. The advantage of the exponential common factor structure for Λ_{it} in equation (2) is that it allows us to distinguish between advances in absorptive capacity that lead to *level* versus *growth* shifts in a country's TFP. To illustrate this, it is convenient to study a Taylor expansion of Λ_{it}

$$\Lambda_{it} = e^{a_{it} + \vartheta_{it} f_t} = (1 + a_{it}) + (1 + a_{it}) \vartheta_{it} f_t + \dots, \tag{4}$$

together with the growth rate of Λ_{it}

$$\Delta \lambda_{it} = \Delta a_{it} + \Delta \vartheta_{it} f_{t-1} + \vartheta_{it} \Delta f_t. \tag{5}$$

In the absence of any changes in a_{it} and ϑ_{it}, the level and growth rate of a country's TFP Λ_{it} are both fixed proportions of the level and growth of global TFP f_t. Equation (4) shows that an increase in a_{it} implies that a country is able to assimilate more of the global technology f_t, while from equation (5) it is clear that this leaves the future TFP growth rate unaffected. Hence, advances in absorptive capacity that lead to a level shift in a country's TFP will be captured by changes in a_{it}. Policy interventions or other efforts that lead to an increase in ϑ_{it} induce a similar TFP level shift but, as apparent from the term $\vartheta_{it} \Delta f_t$ in equation (5), these also imply that the economy's TFP will now grow at a *permanently* higher rate as *more* of the global TFP growth is absorbed.

Further note that the limiting case where countries grow in isolation of the world (as in early endogeneous growth models) is obtained by setting $\vartheta_{it} = 0$ in equation (3) with a_{it} then capturing country-specific growth. Moreover, the neoclassical growth model is also nested in our framework. Here, as Denison

(1967) notes, global knowledge is assumed available to all countries without restrictions, implying that in our framework $\vartheta_{it} = 1$ *in combination with* $a_{it} = a$ (e.g. Mankiw et al., 1992).[5]

2.2 CCE approach to identify unobserved worldwide technology

Worldwide technology f_t in equation (3) is unobserved and can therefore not directly be included as an observed regressor. However, given that f_t is a factor common to all countries, it can be inferred from the cross-sectional dimension of the panel. Factor-augmented regressions have become very popular in the recent panel data literature. A number of alternative estimation procedures are available. The dynamic factor model literature specifies a data generating process for f_t (e.g., a random walk with time-varying drift) and then casts the model in state space form such that f_t can be estimated from the data using the Kalman filter (see Stock and Watson, 2016, for a general discussion). Alternative estimators rely on cross-sectional averages or principal component analysis (see Westerlund and Urban, 2015, for a comparison of these two approaches). Important to note is that, independently of the estimation approach, only the product $\vartheta_{it}f_t$ is identified but not its constituent components: multiplying the loadings ϑ_{it} by a rescaling constant c while dividing the common factor f_t by the same c would leave the product unchanged. This rotation indeterminacy can be solved by imposing normalization constraints on the loadings and/or the factor. In dynamic factor models, a standard normalization is to constrain the scale of the factor f_t by fixing the variance of its innovations to be one. However, while being effective in a model with fixed loadings, time variation brings about challenges to identification as the rescaling term can now be a time-varying sequence c_t rather than a constant c. A further identification problem may arise when separating the idiosyncratic components a_{it} and ϑ_{it}: although these are assumed to be uncorrelated across countries, this is not explicitly imposed by the Kalman filter used to estimate them, such that there is some scope for a_{it} to pick up common technology trends that should in fact be captured by $\vartheta_{it}f_t$.

The above mentioned identification issues can be avoided by using cross-sectional averages to proxy the unobserved common factor f_t. Inspired by the CCE approach of Pesaran (2006), taking cross-sectional averages of the model in equation (3) yields

$$\overline{\lambda}_t = \overline{a}_t + \overline{\vartheta}_t f_t + \overline{\epsilon}_t, \tag{6}$$

where $\overline{\lambda}_t = \frac{1}{N}\sum_{i=1}^{N} y_{it}$ and similarly for the other variables. Solving for f_t

$$f_t = \frac{1}{\overline{\vartheta}_t}\left(\overline{\lambda}_t - \overline{a}_t - \overline{\epsilon}_t\right), \tag{7}$$

---

[5]Since it is widely accepted that TFP level differences across countries are a salient feature of the data (e.g. Islam, 1995) this aspect can be accommodated by setting $a_{it} = a_i$. But this is a 'stylized fact' derived from the data rather than from the neoclassical model and rules out absolute convergence.

and substituting this solution back into equation (3) yields

$$\lambda_{it} = \left(a_{it} - \frac{\vartheta_{it}}{\overline{\vartheta}_t} \overline{a}_t \right) + \frac{\vartheta_{it}}{\overline{\vartheta}_t} \overline{\lambda}_t + \left(\epsilon_{it} - \frac{\vartheta_{it}}{\overline{\vartheta}_t} \overline{\epsilon}_t \right), \tag{8}$$

$$= \alpha_{it} + \theta_{it} \overline{\lambda}_t + \varepsilon_{it}, \tag{9}$$

where $\alpha_{it} = a_{it} - \frac{\vartheta_{it}}{\overline{\vartheta}_t} \overline{a}_t$, $\theta_{it} = \frac{\vartheta_{it}}{\overline{\vartheta}_t}$ and $\varepsilon_{it} = \epsilon_{it} - \frac{\vartheta_{it}}{\overline{\vartheta}_t} \overline{\epsilon}_t$. Given the assumption that ϵ_{it} is a zero-mean white noise term uncorrelated across cross-sectional units, we have that $\overline{\epsilon}_t \xrightarrow{p} 0$ as $N \to \infty$ such that equation (6) implies that the cross-sectional average $\overline{\lambda}_t$ of TFP can be used as an observable proxy for the rescaled and recentered factor $\overline{\vartheta}_t f_t + \overline{a}_t$.

It is easily verified that the implicit normalizations imposed when going from equation (1) to (9) solve the identification issues outlined earlier in this subsection. First, the scale of $\overline{\lambda}_t$, used as a proxy for f_t, is determined as it can be calculated directly from the observed data. Second, the factor loadings θ_{it} are normalized to be one on average across countries in every period, i.e. $\frac{1}{N} \sum_{i=1}^{N} \theta_{it} = \frac{1}{N} \sum_{i=1}^{N} \vartheta_{it} / \overline{\vartheta}_t = 1$ $\forall t$, such that they can no longer be multiplied by a time-varying sequence c_t. Third, the cross-sectional average of a_{it} is normalized to zero in every period, i.e. $\frac{1}{N} \sum_{i=1}^{N} \alpha_{it} = \frac{1}{N} \sum_{i=1}^{N} (a_{it} - \frac{\vartheta_{it}}{\overline{\vartheta}_t} \overline{a}_t) = 0$ $\forall t$, such that it cannot pick up common technology trends.

Note that the cross-sectional average $\overline{\lambda}_t$ of TFP used as a proxy for $\widehat{f}_t$ should not be interpreted as the world TFP frontier but rather as an index of average world technology, with the combination of α_{it} and θ_{it} indicating whether a country operates below or above this average. Nevertheless, the normalized versions α_{it} and θ_{it} can still be used for the main goal of the paper, which is testing the contrasting predictions of alternative growth models about the long-run effects of a country's investment in absorptive capacity development. In particular, the proposition of the multicountry endogenous growth model (as outlined in Section 2.1 above) that the original ϑ_{it} is one implies that also the rescaled θ_{it} is one, while the alternatives that $\vartheta_{it} \neq 1$ or is time-varying will be reflected in θ_{it} as well.

2.3 Modeling and testing for time-varying absorptive capacity

At the heart of our paper are the time-varying parameters α_{it} and θ_{it} measuring a country's capacity to incorporate the world technology into its own production techniques. This time variation implies that equation (9) cannot be estimated using the standard CCE approach of Pesaran (2006). As an alternative, we set up a state space model. We further use a Bayesian model specification search procedure to analyze whether our generalization to a time-varying parameters setting is empirically relevant. If the restrictions $\alpha_{it} = \alpha_i$ and $\theta_{it} = \theta_i$ are valid, our model simplifies to a standard common factor error structure that can be estimated using the conventional CCE approach.

State space model

We complete the model by assuming that the absorptive capacity parameters α_{it} and θ_{it} evolve according to random walk processes

$$\alpha_{it} = \alpha_{it-1} + \psi_{it}^{\alpha}, \qquad\qquad \psi_{it}^{\alpha} \sim \mathcal{N}(0, \varsigma_{\alpha}), \qquad\qquad (10)$$

$$\theta_{it} = \theta_{it-1} + \psi_{it}^{\theta}, \qquad\qquad \psi_{it}^{\theta} \sim \mathcal{N}(0, \varsigma_{\theta}). \qquad\qquad (11)$$

The random walk assumption allows for a very flexible evolution of the parameters over time. The model can then be cast in a state space representation with (9) being the 'observation equation', where for the noise term ε_{it} we assume $\varepsilon_{it} \sim \mathcal{N}\left(0, \sigma_{\varepsilon}^2\right)$, and (10)-(11) the 'state equations' such that the random walk components α_{it} and θ_{it} can be estimated using the Kalman filter.

Bayesian stochastic model specification search

To determine whether the proposed time variation in the parameters α_{it} and θ_{it} is relevant implies that we have to test whether the innovation variances ς_{α} and ς_{θ} in equations (10)-(11) are zero or not. From a classical point of view this is cumbersome as the null hypothesis of zero variance lies on the boundary of the parameter space. We therefore use the stochastic model specification search of Frühwirth-Schnatter and Wagner (2010), generalizing standard Bayesian variable selection to state space models. This involves reparameterizing the state equations (10)-(11) to

$$\alpha_{it} = \alpha_{i0} + \sqrt{\varsigma_{\alpha}}\widetilde{\alpha}_{it}, \qquad \text{with } \widetilde{\alpha}_{it} = \widetilde{\alpha}_{i,t-1} + \widetilde{\psi}_{it}^{\alpha}, \quad \widetilde{\alpha}_{i0} = 0, \quad \widetilde{\psi}_{it}^{\alpha} \sim \mathcal{N}\left(0,1\right), \quad (12)$$

$$\theta_{it} = \theta_{i0} + \sqrt{\varsigma_{\theta}}\widetilde{\theta}_{it}, \qquad \text{with } \widetilde{\theta}_{it} = \widetilde{\theta}_{i,t-1} + \widetilde{\psi}_{it}^{\theta}, \quad \widetilde{\theta}_{i0} = 0, \quad \widetilde{\psi}_{it}^{\theta} \sim \mathcal{N}\left(0,1\right), \quad (13)$$

which splits α_{it} and θ_{it} into the initial values α_{i0} and θ_{i0} and the (possibly) time-varying parts $\sqrt{\varsigma_{\alpha}}\widetilde{\alpha}_{it}$ and $\sqrt{\varsigma_{\theta}}\widetilde{\theta}_{it}$.

This 'non-centered' parameterization has a number of interesting features. First, the signs of both $\sqrt{\varsigma_{\alpha}}$ and $\widetilde{\alpha}_{it}$ can be changed without changing their product, and similarly for $\sqrt{\varsigma_{\theta}}$ and $\widetilde{\theta}_{it}$. This lack of sign identification offers a first piece of information about whether time variation is relevant or not: for truly time-*varying* parameters, the innovation variance ς will be positive resulting in a posterior distribution of $\sqrt{\varsigma}$ that is bimodal with modes $\pm\sqrt{\varsigma}$. For time-*invariant* parameters, ς is zero such that $\sqrt{\varsigma}$ becomes unimodal at zero.

Second, the non-centered parameterization is very useful for model selection as it represents α_{it} and θ_{it} as a superposition of the initial values α_{i0} and θ_{i0} and the time-varying components $\widetilde{\alpha}_{it}$ and $\widetilde{\theta}_{it}$. As a result, in contrast to the centered parameterization in equations (10)-(11), $\widetilde{\alpha}_{it}$ and $\widetilde{\theta}_{it}$ do not degenerate to a static component when the innovation variances are zero. In fact, when for instance $\varsigma_{\alpha} = 0$, then $\sqrt{\varsigma_{\alpha}} = 0$ such that $\widetilde{\alpha}_{it}$ will drop from the model. As suggested by Frühwirth-Schnatter and Wagner

(2010), this allows us to transform the test of whether the variances ς_α and ς_θ are zero or not into a more regular variable selection problem. To this end we introduce two binary indicators δ_α and δ_θ, which are equal to one if the corresponding parameter varies over time and zero otherwise. The resulting parsimonious non-centered specification is given by

$$\lambda_{it} = \alpha_{i0} + \delta_\alpha \sqrt{\varsigma_\alpha} \widetilde{\alpha}_{it} + \left(\theta_{i0} + \delta_\theta \sqrt{\varsigma_\theta} \widetilde{\theta}_{it} \right) \overline{\lambda}_t + \varepsilon_{it}. \tag{14}$$

When $\delta_\alpha = 1$, α_{i0} is the initial value of α_{it} and $\sqrt{\varsigma_\alpha}$ is an unconstrained parameter that is estimated from the data. In contrast, when $\delta_\alpha = 0$ the time-varying part $\widetilde{\alpha}_{it}$ drops out and α_{i0} represents the time-invariant parameter. A similar interpretation holds for δ_θ.

A third important advantage of the non-centered parameterization is that it allows us to replace the standard Inverse Gamma prior on the variance parameters ς_α and ς_θ by a Gaussian prior centered at zero on $\sqrt{\varsigma_\alpha}$ and $\sqrt{\varsigma_\theta}$.[6] Centering the prior distribution at zero is possible as for both $\varsigma = 0$ and $\varsigma > 0$, $\sqrt{\varsigma}$ is symmetric around zero, with the main difference being that in the latter case the posterior distribution is bimodal.

2.4 MCMC algorithm

The state space representation in equations (12)-(14) is a non-linear model for which the standard approach of using the Kalman filter to obtain the time-varying components and Maximum Likelihood to estimate the unknown parameters is inappropriate. We therefore use an MCMC approach to jointly sample the binary indicators $\boldsymbol{\delta} = \{\delta_\alpha, \delta_\theta\}$, the unrestricted elements of the parameter vector $\boldsymbol{\phi} = \{\alpha_{i0}, \theta_{i0}, \sqrt{\varsigma_\alpha}, \sqrt{\varsigma_\theta}, \sigma_\varepsilon^2\}_{i=1}^N$ and the latent state processes $\boldsymbol{s} = \{\{\widetilde{\alpha}_{it}, \widetilde{\theta}_{it}\}_{t=1}^T\}_{i=1}^N$ from the posterior distribution $g(\boldsymbol{\delta}, \boldsymbol{\phi}, \boldsymbol{s}|\boldsymbol{x})$ conditional on the data $\boldsymbol{x} = \{\{\lambda_{it}\}_{t=1}^T\}_{i=1}^N$. This conveniently splits the non-linear estimation problem into a sequence of blocks which are linear conditional on the other blocks. Given a set of starting values, sampling from the various blocks is iterated K times and, after a sufficiently long burn-in period B, the sequence of draws $(B+1, ..., K)$ approximates a sample from $g(\boldsymbol{\delta}, \boldsymbol{\phi}, \boldsymbol{s}|\boldsymbol{x})$. The results reported below will be based on $K = 200,000$ iterations with a burn-in of $B = 20,000$ draws. Following Frühwirth-Schnatter and Wagner (2010), we fix the binary indicators in $\boldsymbol{\delta}$ to be one during the first $10,000$ iterations of the burn-in period to obtain sensible starting values for the unrestricted model before variable selection actually starts. More details on our MCMC algorithm can be found in Appendix A, which includes a detailed description of the different building blocks, an interweaving approach to boost the mixing efficiency, and a simulation exercise tailored to our dataset to investigate the performance of the Bayesian model specification search and the small sample properties of the estimation method.

---

[6]Frühwirth-Schnatter and Wagner (2010) show that compared to using an Inverse Gamma prior for ς, the posterior density of $\sqrt{\varsigma}$ is much less sensitive to the hyperparameters of the Gaussian distribution and, importantly, is not pushed away from zero when $\varsigma = 0$.

3 Data, prior choices and results

3.1 Data

We estimate the empirical specification outlined in Section 2 for a panel of 31 advanced economies using annual data over the period 1953-2014 taken from the Penn World Table (PWT) version 9 (Feenstra et al., 2015). Our sample is made up of 26 current OECD member countries (comprising all current members with the exception of Israel, Turkey and the seven former transition economies), with the addition of Argentina, Brazil, Colombia, Cyprus and Taiwan. At the start of our sample these 31 countries account for over 80% of world GDP, declining to just under half in 2014. Table 1 provides details of the data construction. Real GDP and the real capital stock are in constant 2011 national prices transformed into 2011 US$ values. The capital stock defined by PWT version 9 includes residential structures. Total hours worked is calculated by multiplying the number of persons engaged times the average annual hours worked per person. The income share of capital stock, used as a proxy for β_i, is measured as 1 minus the economy-wide labor share averaged over the sample period.

Figure 1 plots our TFP measure, calculated as $y_{it} - \beta_i x_{it}$, averaged over countries together with its growth rate. In line with the productivity growth patterns documented by Blanchard (2004), Madsen (2008) and van Ark et al. (2008), our average TFP measure shows that the post-war period can be split into three episodes: First, the 1950s and 1960s are a period of high global TFP growth in excess of 2% per annum. Second, the early 1970s herald an era of lower growth with a steep decline in TFP growth. Third, a slight improvement during the 1990s, whereafter global TFP growth nosedives during the Global Financial Crisis (GFC) in 2007/8 and subsequently seems to stabilize around 0%.[7]

Table 1: Data construction and sources

Name	Notation	Construction	Code
Real GDP (in million US$, 2005 values)	Y_{it}	PWT data	rgdpna
Real Capital stock (in million US$, 2005 values)	K_{it}	PWT data	rkna
Number of persons engaged	N_{it}	PWT data	emp
Average annual hours worked by persons engaged	H_{it}	PWT data	avh
Total annual hours worked in the economy	L_{it}	$N_{it} \times H_{it}$	
Income share of labor stock in GDP	ω_{it}	PWT data	labsh
Income share of capital stock in GDP	β_i	$1 - \frac{1}{T}\sum_1^T \omega_{it}$	
Log output per hour worked	y_{it}	$\ln(Y_{it}/L_{it})$	
Log capital per hour worked	k_{it}	$\ln(K_{it}/L_{it})$	
Log TFP	λ_{it}	$y_{it} - \beta_i k_{it}$	

---

[7]This raises the question whether the GFC signals a new era of stagnant global TFP — with the data restrictions on time since the GFC we are unable to address this in the present study.

Figure 1: Cross-country average TFP $\overline{\lambda}_t$ and its growth rate

(a) Level of $\overline{\lambda}_t$ **(b)** Growth rate of $\overline{\lambda}_t$
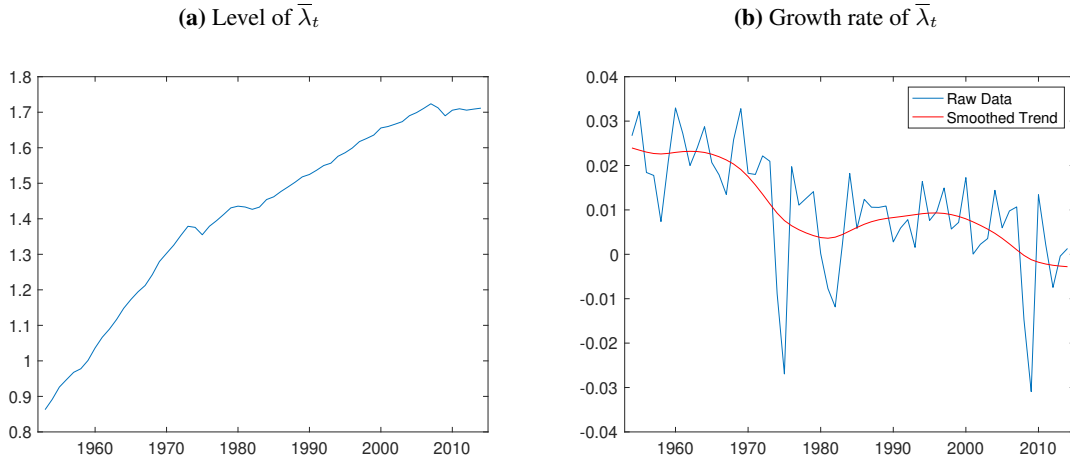
Note: The smoothed trend in average TFP growth was obtained after applying the Hodrick-Prescott filter with smoothing parameter equal to 100.

3.2 Prior choices

For the variance σ_ε^2 of the errors terms ε_{it} we use an uninformative Inverse Gamma prior distribution $IG\left(c_0, C_0\right)$, with the shape c_0 and scale C_0 parameters both set to 0.001. For all other parameters we use a Gaussian prior distribution $N\left(a_0, A_0\sigma_\varepsilon^2\right)$ defined by setting a prior belief a_0 with prior variance $A_0\sigma_\varepsilon^2$. Throughout the estimation procedure we fix A_0 to 100^2 to ensure that posterior results are driven by the information contained in the data and not by the priors.[8] Note that the non-centered parameterization in equation (14) allows us to make use of a Gaussian prior for the standard deviations $\sqrt{\varsigma_\alpha}$ and $\sqrt{\varsigma_\theta}$ since estimating them boils down to a standard linear regression. When sampling the indicators $\boldsymbol{\delta}$ we assign 50% prior probability that the indicators take on the value 1, i.e. $p_0 = 0.5$. For the initial values of the time-varying processes, α_{i0} and θ_{i0}, our prior beliefs are 0 and 1, respectively. This is a natural outcome of the way the CCE approach normalizes α_{it} and θ_{it}, i.e. $\frac{1}{N}\sum_{i=1}^{N}\alpha_{it} = 0$ and $\frac{1}{N}\sum_{i=1}^{N}\theta_{it} = 1$ for all t and hence also for the initial values α_{i0} and θ_{i0}. Our prior belief for $\sqrt{\varsigma_\alpha}$ and $\sqrt{\varsigma_\theta}$ is 0. We center this distribution around zero such that our belief is in accordance with the null hypothesis of our test for whether α_{it} and θ_{it} are fixed over time.

3.3 Empirical results

**Time variation in the absorptive capacity parameters: level versus growth effects**

We start by discussing the results of the stochastic model specification search to analyze whether time variation in the absorptive capacity parameters α_{it} and θ_{it} is a relevant aspect of the model. This enables us to discriminate between the four possible models nested in our setup, i.e. a model where changes in absorptive capacity lead to either growth or level shifts in TFP, a combination of the two, or a model

---

[8]Given that we use uninformative priors, the results reported below are robust to choosing alternative prior beliefs.

without any changes in absorptive capacity. This speaks directly to our distinction between the theoretical models discussed in Section 2.

As a first step, we fix the binary indicators δ_α and δ_θ at 1 to obtain posterior distributions for the unrestricted model where both α_{it} and θ_{it} are allowed to vary over time. When time variation is relevant, this should show up as bimodality in the posterior distribution of the corresponding innovation standard deviation $\sqrt{\varsigma}$. A unimodal distribution centered at zero is expected for time-invariant parameters. Figure 2 plots the posterior distributions of $\sqrt{\varsigma_\alpha}$ and $\sqrt{\varsigma_\theta}$. The results are decisive in that the posterior distribution of $\sqrt{\varsigma_\alpha}$ is bimodal while that of $\sqrt{\varsigma_\theta}$ is perfectly unimodal, offering a strong indication that the data answer to a model where changes in absorptive capacity lead to TFP level rather than growth shifts.

Figure 2: Posterior distributions of $\sqrt{\varsigma_\alpha}$ and $\sqrt{\varsigma_\theta}$

(a) $\sqrt{\varsigma_\alpha}$ — level shift
(b) $\sqrt{\varsigma_\theta}$ — growth shift
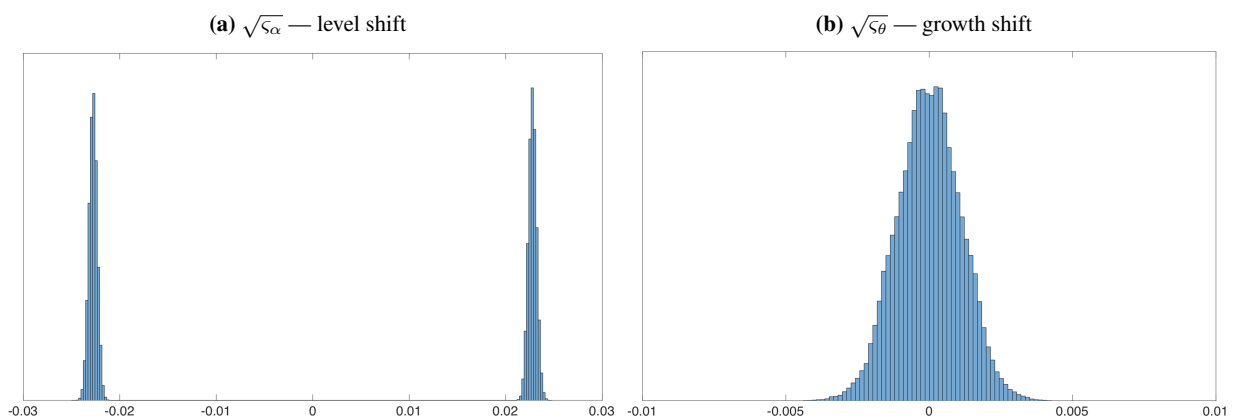
Note: Reported are the posterior distributions of the standard deviations $\sqrt{\varsigma_\alpha}$ and $\sqrt{\varsigma_\theta}$ in equation (14) setting $\delta_\alpha = \delta_\theta = 1$, based on MCMC with $200{,}000$ iterations after a burn-in of $20{,}000$ draws. The integrated autocorrelation time is 1.9 for $|\sqrt{\varsigma_\alpha}|$ and 91.9 for $|\sqrt{\varsigma_\theta}|$.

As a more formal test for time variation, we sample the stochastic binary indicators δ_α and δ_θ together with the other parameters in the model. The upper panel of Table 2 reports the posterior probabilities for the binary indicators being one, calculated as the fraction of MCMC draws for which the stochastic model specification search prefers a model which allows for time variation in the corresponding parameter. It is clear that time variation is important as the test assigns a 100% posterior inclusion probability for $\delta_\alpha = 1$. The probability that $\delta_\theta = 1$, however, is less than 1%. This finding supports our previous conclusion that α_{it} exhibits relevant time variation while θ_{it} is most likely constant over time. Taken together this suggests that in our sample we find evidence *against* the Benhabib and Spiegel (2005) type of model where the long-run growth rate of TFP can be altered using policy interventions. The lower panel of Table 2 reports summary statistics for the posterior distributions of $\sqrt{\varsigma_\alpha}$, $\sqrt{\varsigma_\theta}$ and σ_ε. In line with the model specification search, the median of the posterior distribution of $\sqrt{\varsigma_\theta}$ is very close to zero, while that of $\sqrt{\varsigma_\alpha}$ has no probability mass at zero. Imposing the restriction that $\theta_{it} = \theta_i$ does not change the posterior distributions of $\sqrt{\varsigma_\alpha}$ and σ_ε.

Table 2: Posterior estimation results — model selection on α_{it} and θ_{it}

	Models		
	Unrestricted	Parsimonious ($\theta_{it} = \theta_i$)	Parsimonious — hybrid
Probability of time variation			
α_{it}	1.00		
θ_{it}	0.00		
Parameter summary statistics			
$\left\| \sqrt{\varsigma_\alpha} \right\|$	0.0228	0.0228	0.0228
	(0.0004)	(0.0004)	(0.0004)
$\left\| \sqrt{\varsigma_\theta} \right\|$	0.0005		
	(0.0006)		
σ_ε	0.0058	0.0058	0.0060
	(0.0003)	(0.0003)	(0.0003)

Notes: The unrestricted model allows for time variation in both a_{it} and θ_{it} (binary indicators δ_α and δ_θ sampled in the top panel and set to 1 in the bottom panel); the parsimonious model restricts θ_{it} to be time invariant ($\delta_\theta = 0$) but allows heterogeneity across countries; the parsimonious–hybrid model further sets $\theta_i = 1$ for those countries where the posterior probability that $\theta_i \neq 1$ is smaller than 50%. The priors are as outlined in Section 3.2. The probabilities of time variation are calculated as the fraction of MCMC draws in which the stochastic model specification search prefers a model which allows for time variation in the corresponding parameter. The summary statistics for the parameters are the median and the standard deviation (in parentheses) of their posterior distribution. Results based on MCMC with $200,000$ iterations after a burn-in of $20,000$ draws. The average integrated autocorrelation time (across the parameters of the different models) is 13.0.

In the long-run do all countries grow at the same pace?

Based on the stochastic model specification search, we can conclude that there is relevant time variation in α_{it} but not in θ_{it}. The latter still allows for $\theta_i \neq 1$, whereas an intrinsic property of the multicountry endogenous growth models is that $\theta_i = 1$, such that in the long-run all countries grow at the same pace. Table 3 reports posterior results for θ_i obtained from estimating a parsimonious specification where we set $\delta_\alpha = 1$ and $\delta_\theta = 0$. For most, but not all, countries the 90% highest density interval (HDI) includes 1. In order to test this in a more rigorous way, we can again use the stochastic variable selection approach. To this end we (i) split θ_i into 1 and its deviation ($\theta_i - 1$), and (ii) add a binary indicator $\gamma_{i\theta}$ that equals one when the corresponding variable $(\theta_i - 1)\overline{\lambda}_t$ should be included in the model and zero otherwise. This results in the following specification

$$\lambda_{it} - \overline{\lambda}_t = \alpha_{it} + \gamma_{i\theta}(\theta_i - 1)\overline{\lambda}_t + \varepsilon_{it}, \tag{15}$$

where for $\gamma_{i\theta} = 1$ the deviation of θ_i from one is estimated from the data while $\gamma_{i\theta} = 0$ implies that $\theta_i = 1$.

Table 3 reports posterior inclusion probabilities for the term $(\theta_i - 1)\overline{\lambda}_t$, calculated as the frequency that $\gamma_{i\theta}$ takes on the value of one over the MCMC iterations. For most countries the results are in line with the model of Klenow and Rodríguez-Clare (2005) as deviations of θ_i from one are not found to be a relevant aspect of the model. However, for a number of countries the restriction that $\theta_i = 1$ is not supported by the data. This is most prominently the case for Argentina, Chile, Cyprus, Portugal and Taiwan, for

which the posterior model inclusion probability of $(\theta_i - 1)\overline{\lambda}_t$ clearly exceeds 75%, and to a lesser extent for Australia, Canada, Germany and USA, where the posterior model inclusion probability of $(\theta_i - 1)\overline{\lambda}_t$ is between 50% and 75%. This suggests that over the period 1953-2014 a number of countries in our dataset did have TFP growth that permanently differed from the global evolution. However, many of the aforementioned countries typically are those that have caught up to (Cyprus, Germany, Portugal and Taiwan) or have been caught-up by (Australia, Canada and USA) the global TFP evolution. As far as our sample covers a prolonged period of successful catch-up, this may have resulted in $\theta_i \neq 1$ instead of showing up as time-variation in α_{it}. A longer sample may be needed to rule out this possibility. For completeness, the lower panel of Table 2 also reports results for the posterior distributions of $\sqrt{\varsigma_\alpha}$, and σ_ε in a parsimonious hybrid model where θ_{it} is restricted to 1 only for those countries where the posterior probability that $\theta_i \neq 1$ is smaller than 50%. Compared to the less restricted models in the first two columns, imposing this hybrid restriction does not affect the posterior distributions of $\sqrt{\varsigma_\alpha}$ and σ_ε.

Table 3: Posterior estimation results — model selection on θ_i

	θ_i		$\gamma_{i\theta}$		θ_i		$\gamma_{i\theta}$
Argentina	0.64	(0.15)	0.82	Italy	1.11	(0.15)	0.24
Australia	0.69	(0.15)	0.69	Japan	1.08	(0.15)	0.22
Austria	1.20	(0.15)	0.40	Luxembourg	1.22	(0.15)	0.43
Belgium	1.07	(0.15)	0.23	Mexico	0.79	(0.15)	0.41
Brazil	1.06	(0.15)	0.21	Netherlands	1.09	(0.15)	0.23
Canada	0.69	(0.15)	0.68	New Zealand	0.85	(0.15)	0.29
Chile	0.51	(0.15)	0.98	Norway	1.00	(0.15)	0.21
Colombia	0.77	(0.15)	0.45	Portugal	1.34	(0.15)	0.76
Cyprus	1.34	(0.15)	0.75	South Korea	1.00	(0.15)	0.20
Denmark	0.87	(0.15)	0.29	Spain	1.23	(0.15)	0.47
Finland	1.22	(0.15)	0.43	Sweden	0.75	(0.15)	0.48
France	1.27	(0.15)	0.50	Switzerland	0.81	(0.15)	0.38
Germany	1.27	(0.15)	0.55	Taiwan	1.41	(0.15)	0.91
Greece	1.20	(0.15)	0.39	UK	0.80	(0.15)	0.39
Iceland	0.97	(0.15)	0.20	USA	0.75	(0.15)	0.52
Ireland	1.02	(0.15)	0.21				

Notes: The summary statistics for θ_i are the median and the standard deviation (in parentheses) of the posterior distributions obtained from a parsimonious specification where we restrict $\delta_\alpha = 1$ and $\delta_\theta = 0$. The posterior inclusion probabilities for $\gamma_{i\theta}$ are calculated as the fraction of MCMC draws in which the stochastic model specification search prefers a model with $\gamma_{i\theta} = 1$ (which implies that $\theta_i \neq 1$). Results based on MCMC with $200,000$ iterations after a burn-in of $20,000$ draws. The average integrated autocorrelation time (across the reported parameters) is 16.2.

Figure 3: Posterior results for the absorptive capacity parameter α_{it}
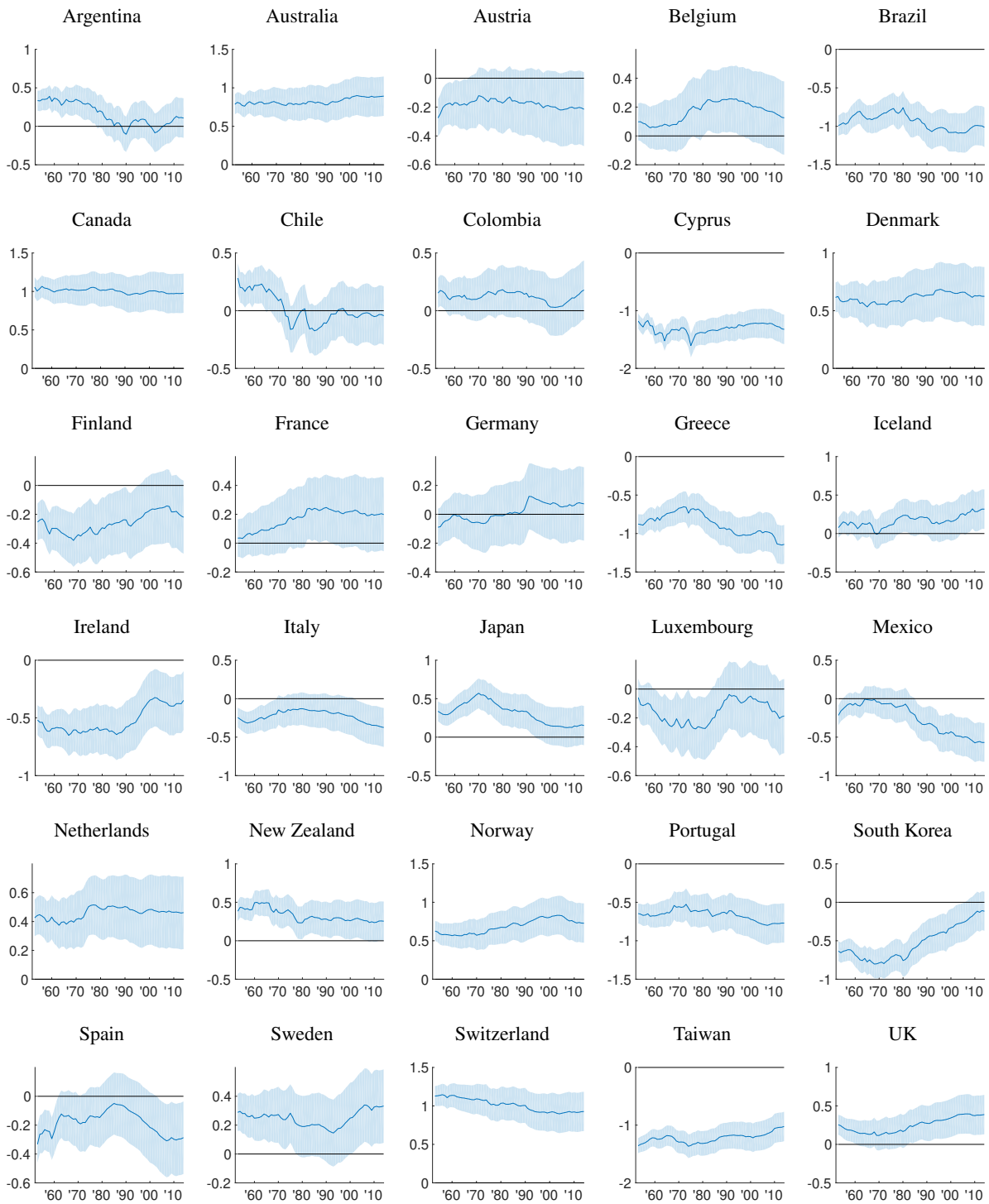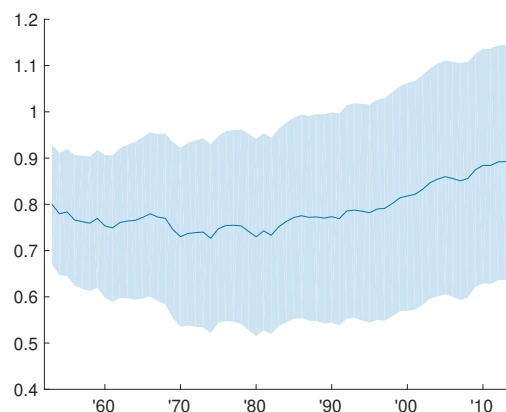

Notes: Reported are the posterior mean and 68% highest density interval (HDI) for a parsimonious model setting $\theta_{it} = \theta_i$. Results for the global technology leader (US) are presented in Figure 4 below. Based on MCMC with $200,000$ iterations after a burn-in of $20,000$ draws. The average integrated autocorrelation time (across the plotted α's) is 1.

Absorptive capacity evolution

Figure 3 presents posterior results for the time-varying absorptive capacity parameter α_{it} in the parsimonious model where $\theta_{it} = \theta_i$. Results for the parsimonious-hybrid model where $\theta_{it} = 1$ for those countries where the posterior probability that $\theta_i \neq 1$ is smaller than 0.5 are presented in Figure **??** in **??**. Looking at the evolution in the absorptive capacity parameter α_{it}, there is substantial variation over time in many countries.[9] A first group, including Finland, France, Germany, Iceland, Norway, South Korea and UK, show an increase in their ability to assimilate foreign knowledge. Ireland and South Korea are clearly catching up with the rest since they started off well below average absorptive capacity in 1953. The opposite evolution can be observed for a second group, consisting of Argentina, Chile, Greece, Japan, Mexico and Switzerland. Other countries show either a modest increase or decrease, with Australia, Austria, Denmark, France and the Netherlands showing little or no structural movement in α_{it}. The seemingly 'static' nature of the latter group of countries is however somewhat misleading, as would be the same verdict for the global technology leader, the United States, which we chart in Figure 4: recall that the absorptive capacity evolution depicted here is a *relative* index, such that these countries can be highlighted as having kept up a very strong absorptive capacity performance, which in the case of Austria, Belgium, and France, among others, was *consistently on par* or in the case of Australia, Canada, Denmark, Sweden, Norway, and the US *even outpacing* the global developments over this time period. The US have strengthened their position as global technology leader after the global catch-up of the 1950s and 60s and the setbacks of the oil crises in the 1970s. The 2007/8 GFC is clearly marked, though the trajectory in its aftermath is upward (albeit with reference to zero global TFP growth: see Figure 1).

Figure 4: Posterior results for the absorptive capacity parameter α_{it} — United States

Notes: Reported is the posterior mean and 68% highest density interval (HDI) for a parsimonious model setting $\theta_{it} = \theta_i$. Based on MCMC with $200,000$ iterations after a burn-in of $20,000$ draws. The average integrated autocorrelation time (across the plotted α's) is 1.

---

[9]In **??** we for instance pick a number of economies on the basis of their diverging paths relative to the global frontier, and describe their policy evolution in greater detail, highlighting the correspondence with our estimated absorptive capacity evolution.

4 Linking the absorptive capacity evolution to economic policy

While the dynamics and evolution of the absorptive capacity parameters are of great interest in and of themselves, it is particularly meaningful to determine country-specific drivers of absorptive capacity. In this section, we therefore extend the empirical specification presented in Section 2 to allow for *observed* country-specific variables to affect absorptive capacity. We will focus on a set of indicators highlighted in the recent Schumpeterian growth literature dominating the current debate on policy for economic growth: aspects of financial development, human capital, competition policy and the R&D knowledge stock. This addition of covariates complicates matters substantially, though as our specification tests reveal our parameters of interest are identified since none of these candidate determinants of absorptive capacity turn out to have time-*variant* parameters.

4.1 Extended empirical specification

We specify the data generating processes of the absorptive capacity parameters a_{it} and ϑ_{it} in equation (3) as

$$a_{it} = a_{it}^* + x_{it}\beta_\alpha, \tag{16}$$

$$\vartheta_{it} = \vartheta_{it}^* + x_{it}\beta_\theta, \tag{17}$$

with x_{it} a K-dimensional vector of observed variables demeaned over time and over cross-sectional units.[10] This more general specification maintains the flexibility of our baseline approach by allowing a_{it} and ϑ_{it} to be driven by the stochastic components a_{it}^* and ϑ_{it}^*, such that *absorptive capacity can still be driven by determinants other than those included* in x_{it}. We allow x_{it} to enter *both* absorptive capacity parameters a_{it} and ϑ_{it} to test whether they are related to shifts in the level or in the growth rate of TFP (see discussion in Section 2.1 above).

Substituting equations (16)-(17) in (3) yields

$$\lambda_{it} = a_{it}^* + x_{it}\beta_\alpha + \left(\vartheta_{it}^* + x_{it}\beta_\theta\right)f_t + \epsilon_{it}. \tag{18}$$

Equivalent to Section 2.2, we use the CCE approach to identify f_t. This involves taking cross-sectional averages of (18), which results in

$$\overline{\lambda}_t = \overline{a}_t^* + \overline{\vartheta}_t^* f_t + \overline{\epsilon}_t, \tag{19}$$

---

[10]Note that a_{it}^* and ϑ_{it}^* in equations (16)-(17) wipe out any individual and time fixed effects from the original data X_{it}. Hence, defining $x_{k,it} = X_{k,it} - \overline{X}_{k,it}$, with $\overline{X}_{k,it} = \frac{1}{N}\sum_{i=1}^N X_{k,it} + \frac{1}{T}\sum_{t=1}^T X_{k,it} - \frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T X_{k,it}$, to be the cross-sectionally and time demeaned version of the kth variable in X_{it}, implies that $\overline{X}_{it}\beta_\alpha$ and $\overline{X}_{it}\beta_\theta$ are absorbed in a_{it}^* and ϑ_{it}^*, respectively.

give that $\overline{x}_t = 0$ by construction. Solving for f_t

$$f_t = \frac{1}{\vartheta_t^*}\left(\overline{\lambda}_t - \overline{a}_t^* - \overline{\epsilon}_t\right),\tag{20}$$

and substituting this solution back into equation (18) yields

$$\lambda_{it} = \alpha_{it}^* + \sum_{k=1}^{K} \beta_{kt} x_{k,it} + \theta_{it}^* \overline{\lambda}_t + \varepsilon_{it},\tag{21}$$

$$\text{where} \qquad \beta_{kt} = \beta_{k\alpha} + \beta_{k\theta}(\overline{\lambda}_t - \overline{a}_t^*)/\overline{\vartheta}_t^*.\tag{22}$$

Further, $\alpha_{it}^* = a_{it}^* - \vartheta_{it}^* \overline{a}_t^*/\overline{\vartheta}_t^*$, $\theta_{it}^* = \vartheta_{it}^*/\overline{\vartheta}_t^*$ and $\varepsilon_{it} = \epsilon_{it} - \frac{\vartheta_{it}^* + x_{it}\beta_\theta}{\overline{\vartheta}_t^*}\overline{\epsilon}_t \rightarrow \epsilon_{it}$ as $N \rightarrow \infty$. Equation (21) shows that an observed variable $x_{k,it}$ that potentially affects absorptive capacity can simply be added to the baseline specification. As equation (22) indicates, $x_{k,it}$ will only have a time-varying impact β_{kt} when it has growth effects. In the absence of growth effects, β_{kt} reduces to the level effect $\beta_{k\alpha}$.

In principle, we can estimate an unrestricted version of equation (21) by allowing for separate time-varying coefficients on each of the observed variables $x_{k,it}$. However, this ignores that the time variation in β_{kt} is driven by a common component $(\overline{\lambda}_t - \overline{a}_t^*)/\overline{\vartheta}_t^*$. To take this restriction into account, we reparameterize equation (21) to

$$\lambda_{it} = \alpha_{it}^* + \beta_{jt} x_{j,it} + \sum_{k=1,k\neq j}^{K} (\beta_{k\alpha}^* + \beta_{jt}\beta_{k\theta}^*)x_{k,it} + \theta_{it}^* \overline{\lambda}_t + \varepsilon_{it},\tag{23}$$

such that $\beta_{jt} = \beta_{j\alpha} + \beta_{j\theta}(\overline{\lambda}_t - \overline{a}_t^*)/\overline{\vartheta}_t^*$, $\beta_{k\alpha}^* = \beta_{k\alpha} - \beta_{j\alpha}\beta_{k\theta}/\beta_{j\theta}$ and $\beta_{k\theta}^* = \beta_{k\theta}/\beta_{j\theta}$.

Note that $\beta_{k\alpha}$ and $\beta_{k\theta}$ are not separately identified from estimates of β_{jt}, $\beta_{k\alpha}^*$ and $\beta_{k\theta}^*$ obtained from equation (23).[11] Nevertheless, our primary question of interest whether observable characteristics such as financial development or human capital have growth effects ($\beta_{k\theta} \neq 0$) or not ($\beta_{k\theta} = 0$) can still be tested using equation (23). First, $\beta_{j\theta} = 0$ implies that $\beta_{jt} = \beta_{j\alpha} + \beta_{j\theta}(\overline{\lambda}_t - \overline{a}_t^*/\overline{\vartheta}_t^*)$ reduces to $\beta_{j\alpha}$. Thus testing for $\beta_{j\theta} = 0$ can be done by testing for time variation in β_{jt}. Second, conditional on β_{jt}, the rescaled parameters $\beta_{k\alpha}^*$ and $\beta_{k\theta}^*$ are identified.[12] Although this does not allow us to separately identify the original $\beta_{k\alpha}$ and $\beta_{k\theta}$, testing whether $\beta_{k\theta} = 0$ can be done by testing $\beta_{k\theta}^* = 0$.

Further note that if $\beta_{k\theta} = 0$ ($\forall k$), as we will come to find below, then we can estimate the following restricted version of equation (23)

$$\lambda_{it} = \alpha_{it}^* + x_{it}\beta_\alpha + \theta_{it}^* \overline{\lambda}_t + \varepsilon_{it},\tag{24}$$

in which β_α *is* identified.

---

[11]In particular, $\beta_{j\alpha}$ and $\beta_{j\theta}$ cannot be calculated from an estimate of β_{jt} because $\overline{a}_t^*/\overline{\vartheta}_t^*$ is unobserved. This in turn implies that $\beta_{k\alpha}$ and $\beta_{k\theta}$ (for $k \neq j$) are not identified.

[12]Note that choosing a normalizing variable $x_{j,it}$ for which $\beta_{j\theta} = 0$ implies that $\beta_{k\theta}^* = \beta_{k\theta}/\beta_{j\theta} = \beta_{k\theta}/0$ is not identified. When estimating the model we will therefore consider alternative choices of normalizing variable to check robustness.

To complete the model, similarly to the baseline specification, we assume that the unobserved stochastic parts of the absorptive capacity parameters[13] α_{it}^* and θ_{it}^* as well as the time varying parameter β_{jt} behave as random walk processes

$$\alpha_{it}^* = \alpha_{it-1}^* + \psi_{it}^\alpha, \qquad\qquad \psi_{it}^\alpha \sim \mathcal{N}(0, \varsigma_\alpha), \qquad (25)$$

$$\theta_{it}^* = \theta_{it-1}^* + \psi_{it}^\theta, \qquad\qquad \psi_{it}^\theta \sim \mathcal{N}(0, \varsigma_\theta), \qquad (26)$$

$$\beta_{jt} = \beta_{jt-1} + \psi_{jt}^\beta, \qquad\qquad \psi_{jt}^\beta \sim \mathcal{N}(0, \varsigma_\beta). \qquad (27)$$

Testing whether α_{it}^*, θ_{it}^* and β_{jt} are constant or vary over time can again be done using the Bayesian model specification search as outlined in Section 2.3.

4.2   Potential drivers of absorptive capacity

In the following, we put forward a set of policy variables highlighted in the recent Schumpeterian growth literature to potentially affect absorptive capacity (financial development, human capital, competition policy and the R&D knowledge stock) and discuss the data sources. Due to data availability the analysis for this extended specification is carried out on a reduced sample of 21 OECD countries over the 1970-2009 time period.[14]

Financial Development

A large branch of the empirical literature has successfully documented a positive link between a well-developed financial sector and long-run economic growth through capital accumulation and technological progress. In the absence of financial intermediaries, informational asymmetries, transaction costs and liquidity risk can impede an the optimal allocation of capital, such that innovative projects with potentially high returns struggle to find financing (see Levine, 1997, for an in-depth discussion). Well-functioning banks are able to screen new projects at lower costs and can better diversify risk, making it easier to fund those start-ups with the best chance of employing innovative products and production processes. This in turn stimulates technological progress. Theoretical evidence for a positive link between financial development and technological progress can be found *inter alia* in the endogenous growth models of De la Fuente and Marín (1996) and more recently Laeven et al. (2015), with the direct link to technology adoption modeled in Cole et al. (2016). Empirically, King and Levine (1993) confirm that financial services enhance growth by both fostering capital formation and improving the efficiency of that capital stock. Numerous studies by Ross Levine and Thorsten Beck (most notably Levine et al., 2000) point to the positive association between financial innovation on the one hand, and capital allocation efficiency

---

[13]The model in (24) is restricted with respect to the growth effects of observed policy variables only. The level or growth effects of unobserved determinants of absorptive capacity can be controlled by restricting α_{it}^* and θ_{it}^* as before.

[14]Compared with our results in Section 3 the following countries drop out of the sample: Argentina, Brazil, Chile, Colombia, Cyprus, Iceland, Luxembourg, Mexico, South Korea, and Taiwan.

and economic growth on the other.[15] More recently, in the aftermath of the 2007/8 Global Financial Crisis, concerns over 'too much finance' have taken the finance-growth literature into a new direction. In an article of the same name, Arcand et al. (2015) demonstrate that in countries with very large financial sectors, as proxied by private credit/GDP, there is no positive correlation between financial depth and economic growth.

As a proxy for financial development, we use data on total market-value bank credit to the private, non-financial sector, expressed in percent of GDP, from the *Bank of International Settlements* (BIS). This series is adjusted for structural breaks and we imputed missing values splicing the private credit series from the July 2018 update to Beck et al. (2000). In order to allow for the potential of a 'vanishing effect' of financial depth, we include levels and squared terms of the log-transformed debt-to-GDP ratio (Arcand et al., 2015).

Human capital

The study of human capital in relation to economic growth and development has long suffered from a failure to distinguish between the types of knowledge/education 'appropriate' at different levels of development — e.g. the Bils and Klenow (2000) 'puzzle' of comparatively low importance of education for growth; or Prichett's seminal work 'Where has all the education gone?' (Pritchett, 2001). The existing literature on knowledge spillovers typically assumes a human capital angle to absorptive capacity, though studies disagree on the significance of levels versus growth of 'skills' (e.g. Keller, 1996; Kneller and Stevens, 2006).

Our data are taken from the Penn World Table (Feenstra et al., 2015, v.9.1): the annual human capital index combines average years of schooling data from Barro and Lee (2010) and Cohen and Leker (2014) with returns to education estimates at primary, secondary and tertiary level (Psacharopoulos, 1994), hence emphasizing the differences between different levels of average educational attainment.

Regulation

Much of the recent literature on innovation and growth has worked towards solving the often contradictory theoretical and empirical results on the role of competition by taking a more differentiated view of 'pre-innovation' and 'post-innovation' rents (Aghion and Griffith, 2005). The well-known inverted-U shape result of Aghion et al. (2005) for the relationship between competition and sectoral growth is the result of a (positive) escape competition and a (negative) rent-dissipation effect, with the relative magnitudes determined by the technological characteristics of the sector.

Existing empirical work on product market regulation has emphasized the transformative force of competitive conditions on productivity, innovation, and technology diffusion (Conway et al., 2007; Aghion,

---

[15]Beck et al. (2016) further show that financial innovation is linked to a higher appetite for risk, making bank profits more volatile, thus leading to higher losses when banking crises occur. The net effect of financial intermediation, however, is positive.

2016; Cette et al., 2017). Similarly, policy shifts towards greater labor market flexibility are suggested to accelerate productivity and innovation in advanced countries, perhaps *in interaction* with liberalization of product market regulations (Crafts, 2006; Aghion et al., 2009; Aghion, 2016).

Standard measures of competition policy, such as OECD data for labor and product market regulation, are only available from 1990 onward and in 5-year intervals from 1998, respectively (OECD, 2014), which would substantially cut into our sample. Instead, we adopt the regulation measure from the Fraser Institute *Economic Freedom of the World* panel dataset, which combines measures for regulation of the labor and credit markets and for business regulations.[16] These series are available annually from 2000 and every five years prior to that from 1970 — we linearly interpolate the earlier sample years.

R&D

A large theoretical literature attributes a starring role to R&D in generating productivity gains and long-run development owing to the generation of knowledge spillovers (Grossman and Helpman, 1991). This difference between the private and social returns to R&D investments has been used to motivate policy interventions such as tax credits or R&D subsidies: in the famous words of Zvi Griliches (1979), R&D expenditure represents "one of the few variables which public policy can affect in the future" (115). The investigation of domestic R&D stock and weighted foreign R&D stock in their impact on TFP is a popular empirical exercise since the seminal study by Coe and Helpmann (1995), with more recent iterations including Coe et al. (2009) and Ertur and Musolesi (2017). An alternative empirical setup focuses on domestic R&D as a measure for absorptive capacity (Griffith et al., 2003, 2004, the 'second face' of R&D): here, the international technology transfer in the form of knowledge spillovers can only take place if firms or countries have actively developed their ability to assimilate and employ external knowledge (Mancusi, 2008), with sufficient prior experience of 'own'-R&D a prerequisite (an idea going back to Cohen and Levinthal, 1989).

We follow the standard practice in the literature in adopting a country-level measure of business enterprise R&D (BERD) expenditure expressed in terms of GDP share from Madsen and Ang (2016), which we transform into a real US$ R&D expenditure series using appropriate GDP data, and accumulate into stocks via the perpetual inventory method (the Appendix of Coe et al., 2009, describes the detailed steps including base year values), adopting a 5% depreciation rate. The resulting R&D stocks are expressed in per hour worked and log-transformed.

4.3 Empirical results: extended specification

Our analysis proceeds in two steps. First, we use the unrestricted specification (23) to test whether policy variables and investments in absorptive capacity — both observed and unobserved — induce growth

---

[16]Note that only the aggregate but not the three constituent measures of the regulation index are available.

22

effects or merely affect the level of TFP. Second, based on the insights from this model specification search, we then estimate the restricted specification (24). We discuss these elements in turn below. The priors are as outlined in Section 3.2, with an uninformative $N(0, 100^2\sigma_\varepsilon^2)$ prior for the additional parameters $\sqrt{\varsigma_\beta}$, β_α^* and β_θ^*.

Results: unrestricted extended model

The top panel of Table 4 presents results for the Bayesian model specification search on the unrestricted extended model, in turn using each of the considered observed policy variables to normalize the model (see footnote 12 for details). The results confirm our previous findings of time variation in α and time invariance for θ — recall that these capture absorptive capacity other than that driven by the policy variables. First, α_{it}^* is found to be time varying with probability one while the probability that θ_{it}^* is time varying is zero. Second, none of the included observed variables have clear growth effects: the probability that β_{jt} is time-varying is well below 0.5 for all variables. This implies that $\beta_{k\theta} = 0 \ \forall k$ and confirms the result of the baseline model that policy interventions — both unobserved and observed — can affect the *level* of TFP, but not the *growth* rate. This finding is reiterated by the magnitudes and precision of the absolute values of $\sqrt{\varsigma_\alpha}$, $\sqrt{\varsigma_\theta}$ and $\sqrt{\varsigma_\beta}$ reported in the bottom panel of Table 4.

Results: restricted model

Given the probabilities of time variation presented in Table 4, indicating that θ_{it}^* is time invariant and $\beta_{k\theta} = 0 \ \forall k$ such that both unobserved and observed policy variables only have level effects, we estimate the restricted version (24) of the extended model additionally setting $\theta_{it}^* = \theta_i^*$. The estimated level effects β_α of our set of *observed* policy variables are presented in Table 5.[17]

Many existing empirical studies on economic growth emphasize Schumpeter's well-known argument that financial intermediaries foster growth through innovation by mobilizing savings, managing risk, and facilitating transactions (King and Levine, 1993). When financial development is added to TFP or innovation diffusion regressions, however, evidence over and above enabling R&D investment is mixed (Madsen et al., 2010; Comin and Nanda, 2019). In our results for 21 OECD countries, financial development has a strong positive effect on absorptive capacity, yet is subject to diminishing returns, echoing the 'too much finance' argument in recent work on economic growth by Arcand et al. (2015).[18] Note that like these authors we uncover this nonlinearity despite the sample period for this extended model ending almost immediately after the 2007/8 Global Financial Crisis.

---

[17]Although it is not possible to contribute the explanatory power of the model in equation (24) to its constituent components, an indication of the relative importance of the observed variables versus the latent components is that the variance in $x_{it}\widehat{\beta}$ is 11% of the variance in $\widehat{\alpha}_{it}^* + \widehat{\theta}_{it}^*\overline{\lambda}_t$.

[18]In our model marginal returns turn negative at around 53% of credit/GDP. Although comparison with Arcand et al. (2015) is fraught with difficulty due to differences in sample and, especially, methodology, it is safe to say that our estimate is substantially lower than theirs. Note, however, that the marginal returns implied by our parameter estimates suggest a shallow curvature beyond this threshold.

Table 4: Results for the extended empirical specification — model specification search

	Normalizing variable $x_{j,it}$				
	CRE	CRE^2	REG	HC	R&D
Probability of time variation					
α_{it}^*	1.00	1.00	1.00	1.00	1.00
θ_{it}^*	0.00	0.00	0.00	0.00	0.00
β_{jt}	0.08	0.00	0.02	0.40	0.06
Parameter summary statistics					
$\left\|\sqrt{\varsigma_\alpha}\right\|$	0.0136	0.0137	0.0137	0.0135	0.0137
	(0.0006)	(0.0006)	(0.0005)	(0.0006)	(0.0006)
$\left\|\sqrt{\varsigma_\theta}\right\|$	0.0010	0.0010	0.0010	0.0011	0.0011
	(0.0009)	(0.0010)	(0.0009)	(0.0009)	(0.0009)
$\left\|\sqrt{\varsigma_\beta}\right\|$	0.0049	0.0008	0.0013	0.0205	0.0021
	(0.0351)	(0.0005)	(0.0011)	(0.0248)	(0.0028)
σ_ε	0.0054	0.0054	0.0054	0.0053	0.0054
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)

Notes: Reported are results relevant for model selection obtained from estimating the unrestricted extended specification in equation (23). The policy variables included in the extended model are financial development (CRE) in levels and squared terms, regulation (REG), human capital (HC) and R&D stock (R&D) — *all* of these are included in *all* models. Each column reports results for a different choice of normalizing variable $x_{j,it}$ used to test for growth effects. The priors are as outlined in Section 3.2, with an uninformative $N(0, 100^2\sigma_\varepsilon^2)$ prior for the additional parameters $\sqrt{\varsigma_\beta}$, β_α^* and β_θ^*. The reported probabilities of time variation are calculated as the fraction of MCMC draws in which the stochastic model specification search prefers a model which allows for time variation in the corresponding parameter (obtained by sampling the binary indicators). The summary statistics for the parameters are the median and the standard deviation (in parentheses) of their posterior distribution (obtained by setting the binary indicators to 1). Results based on MCMC with $200,000$ iterations after a burn-in of $20,000$ draws. The average integrated autocorrelation time (across the sampled parameters in the different models) is 47.4.

The results for the regulation index suggest no effect on absorptive capacity — this may be due to the constituent components of this index (regulations of the labor and credit markets, business regulation) having heterogeneous effects which wash out in the aggregate regulation index.

Previous work by Griffith et al. (2003, 2004) emphasized a role for human capital in absorptive capacity evolution and hence successful technology transfer in the 1970s and 1980s, though in other studies using data up to a more recent date the effect was not significant (Madsen et al., 2010). In our preferred specification the human capital effect is positive albeit very small and statistically insignificant. Figure **??** in the appendix shows that for this sample of advanced economies the evolution of human capital was linear for virtually every country, which leaves little variance — and hence power — after removing individual and time fixed effects. The study by Ertur and Musolesi (2017), which like ours accounts for cross-section dependence, similarly finds no significant effect of human capital once (domestic and foreign) R&D stock is included in the model.

As the latter qualification signals, the 'second face' of R&D as a means to foster absorptive capacity is clearly borne out by our empirical results. Virtually all of the existing literature on knowledge spillovers and absorptive capacity has found strong evidence in this regard (e.g. Griffith et al., 2004; Madsen et al.,

2010; Ertur and Musolesi, 2017), though our overall findings reiterate that a focus on R&D exclusively is not warranted.

Table 5: Results for the extended empirical specification — restricted version with $\theta_{it}^* = \theta_i^*$ and $\beta_{k\theta} = 0$ ($\forall k$)

| | Level effect β_α of observed policy variables | | | | | | |
CRE	CRE^2	REG	HC	R&D	$|\sqrt{\varsigma_\alpha}|$	σ_ε
0.5145	-0.0648	-0.0051	0.0031	0.1008	0.0137	0.0056
(0.1108)	(0.0135)	(0.0047)	(0.0740)	(0.0235)	(0.0005)	(0.0003)

Notes: The reported results are obtained from estimating the restricted extended specification in equation (24) additionally imposing that $\theta_{it}^* = \theta_i^*$ is time invariant. The policy variables included in the extended model are financial development (CRE) in levels and squared terms, regulation (REG), human capital (HC) and R&D stock (R&D). The priors are as outlined in Section 3.2, with an uninformative $N(0, 100^2\sigma_\varepsilon^2)$ prior for the additional parameters in β_α. Reported are the median and the standard deviation (in parentheses) of the posterior distribution of the parameters. Results bases on MCMC with $200,000$ iterations after a burn-in of $20,000$ draws. The average integrated autocorrelation time (across the reported sampled parameters) is 2.5.

5 Concluding Remarks

Barriers to technology diffusion potentially have dramatic implications for income and welfare: one counterfactual estimate suggests that if countries did not *share their ideas*, world GDP would be a mere 6% of its current level (Klenow and Rodríguez-Clare, 2005, emphasis added). In the spirit of our own investigation we can perhaps rephrase this statement as follows: if countries did not *absorb the available ideas*, world GDP would be a mere 6% of its current level. This statement raises a number of important questions: what is the nature of this absorption process? Do investments in absorptive capacity have perpetual growth effects or do these benefits peter out after a number of years leaving only a level shift? In this paper we have introduced an empirical methodology which provides a simple but powerful tool to help answering these questions. Our contributions relate to (i) the econometric literature in form of an extension of the Pesaran (2006) common correlated effects (CCE) estimators to a setup where factor loadings are allowed to differ over time, a characteristic we test for as part of our implementation; and to (ii) the empirical literature on growth and productivity which to date has operationalized absorptive capacity by adopting proxies such as R&D investments or human capital, while further specifying explicit channels such as trade, FDI, or migration, through which global technology can transfer to individual countries. Our implementation encompasses all of these channels and proxies, but is far less demanding in terms of empirical data requirements to yield country- and time-specific absorptive capacity indices for policy analysis.

Estimating our benchmark model using a panel of 31 advanced economies covering 1953-2014, we gain a number of important insights into the nature of knowledge evolution and the patterns of knowledge diffusion across this group of countries. First, we establish that time variation in absorptive capacity matters, particularly so in a group of high-growth late developers including Ireland, South Korea and Taiwan. Thus, successful development equals successfully improving absorptive capacity. Second, we

establish that the growth boost from improvements in absorptive capacity is a one-off and does not extend into perpetuity: absorptive capacity growth (and implicitly policy aimed to foster this growth) has TFP *level* effects but not growth effects, a finding in line with theoretical models presented in Klenow and Rodríguez-Clare (2005), Howitt (2000), and others. Our third finding relates to the country-specific evolution of absorptive capacity, which can be squared with policy levers to promote innovation and technical change, as advocated by Schumpeterian growth theory. We estimate an extended model using a panel of 21 countries for which we have indicators of financial development, R&D investment, human capital and competition policy/regulation over the period 1970-2009. In line with the baseline model, none of these variables induces permanent growth effects. In our results financial development has a strong positive effect on absorptive capacity, yet is subject to diminishing returns, echoing the 'too much finance' debate in the growth literature (Arcand et al., 2015). Moreover, R&D knowledge stock is found to contribute substantially to our measure of absorptive capacity.

The analysis in this study merely represents a starting point. We can further expand the sample of countries to move away from a focus on countries at the technology frontier and toward a study of the current 'laggards' of economic development: the analysis of absorptive capacity evolution in low- and middle-income countries can provide important insights into the differential policy implications at different levels of development. Especially in low-income countries investment in R&D is almost negligible and the estimated absorptive capacity indices enable us to identify successful countries and/or time periods which in turn can help point to suitable development policy (Cirera and Maloney, 2017). The analysis of advanced countries as in the present study but over a much longer time horizon can help refine our tests of endogenous growth models (Madsen, 2007). Last but not least, the analysis could move away from aggregate economy data and embrace the rich sector-level data in manufacturing for advanced economies (Griffith et al., 2004; Eberhardt et al., 2013), and in agriculture for developing and emerging economies (Fuglie, 2017).

Our empirical results have implications for the use of the Pesaran (2006) CCE estimator in cross-country productivity analysis. The finding of time-varying intercepts combined with time-invariant factor loadings implies that the standard CCE approach is applicable provided that all relevant drivers of absorptive capacity are included as covariates. Alternatively, our methodology allows us to introduce measured inputs in the innovation process (such as R&D stocks) alongside the current time-varying intercept capturing other intangible aspects of productivity and development (Ertur and Musolesi, 2017) and the common factor error structure capturing global technology.

Acknowledgments

References

Abramovitz, M. (1956). *Resource and Output Trends in the United States Since 1870.* National Bureau of Economic Research, Inc.

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of Economic History*, 46(2):385–406.

Acharya, R. C. (2016). ICT use and total factor productivity growth: intangible capital or productive externalities? *Oxford Economic Papers*, 68(1):16–39.

Acharya, R. C. and Keller, W. (2009). Technology transfer through imports. *Canadian Journal of Economics*, 42(4):1411–1448.

Aghion, P. (2016). Competitiveness and growth policy design. In Keuschnigg, C., editor, *Moving to the Innovation Frontier*, chapter 1, pages 5–29. CEPR.

Aghion, P., Askenazy, P., Bourlès, R., Cette, G., and Dromel, N. (2009). Education, market rigidities and growth. *Economics Letters*, 102(1):62–65.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation: An inverted-u relationship. *The Quarterly Journal of Economics*, 120(2):701–728.

Aghion, P. and Griffith, R. (2005). *Competition and Growth: Reconciling Theory and Evidence*. Cambridge: MIT Press.

Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*. MIT Press.

Aghion, P. and Jaravel, X. (2015). Knowledge spillovers, innovation and growth. *The Economic Journal*, 125(583):533–573.

Alfaro, L., Kalemli-Ozcan, S., and Volosovych, V. (2008). Why Doesn't Capital Flow from Rich to Poor Countries? An Empirical Investigation. *The Review of Economics and Statistics*, 90(2):347–368.

Arcand, J. L., Berkes, E., and Panizza, U. (2015). Too much finance? *Journal of Economic Growth*, 20(2):105–148.

Bahar, D., Hausmann, R., and Hidalgo, C. A. (2014). Neighbors and the evolution of the comparative advantage of nations: Evidence of international knowledge diffusion? *Journal of International Economics*, 92(1):111–123.

Barro, R. J. and Lee, J.-W. (2010). A New Data Set of Educational Attainment in the World, 1950-2010. NBER Working Papers 15902.

Beck, T., Chen, T., Lin, C., and Song, F. M. (2016). Financial Innovation: The Bright and the Dark Sides. *Journal of Banking & Finance*, 72:28–51.

Beck, T., Demirgüç-Kunt, A., and Levine, R. (2000). A new database on the structure and development of the financial sector. *The World Bank Economic Review*, 14(3):597–605.

Benhabib, J. and Spiegel, M. M. (1994). The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34(2):143–173.

Benhabib, J. and Spiegel, M. M. (2005). Human Capital and Technology Diffusion. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, chapter 13, pages 935–966. Elsevier.

Bils, M. and Klenow, P. J. (2000). Does schooling cause growth? *American Economic Review*, 90(5):1160–1183.

Blanchard, O. (2004). The Economic Future of Europe. *Journal of Economic Perspectives*, 18(4):3–26.

Bloom, N., Lemos, R., Sadun, R., Scur, D., and Van Reenen, J. (2016). International data on measuring management practices. *American Economic Review: Papers & Proceedings*, 106(5):152–156.

Bosetti, V., Cattaneo, C., and Verdolini, E. (2015). Migration of skilled workers and innovation: A European Perspective. *Journal of International Economics*, 96(2):311–322.

Cette, G., Lopez, J., and Mairesse, J. (2017). Upstream product market regulations, ict, r&d and productivity. *Review of Income and Wealth*, 63:S68–S89.

Cirera, X. and Maloney, W. F. (2017). *The Innovation Paradox: Developing-Country Capabilities and the Unrealized Promise of Technological Catch-Up*. World Bank Publications.

Coe, D. T., Helpman, E., and Hoffmaister, A. W. (2009). International R&D Spillovers and Institutions. *European Economic Review*, 53(7):723–741.

Coe, D. T. and Helpmann, E. (1995). International R&D Spillovers. *European Economic Review*, 39(5):859–887.

Cohen, D. and Leker, L. (2014). Health and education: Another look with the proper data. CEPR Discussion Paper Series 9940.

Cohen, W. M. and Levinthal, D. A. (1989). Innovation and Learning: The Two Faces of R&D. *The Economic Journal*, 99(397):569–596.

Cole, H. L., Greenwood, J., and Sanchez, J. M. (2016). Why doesn't technology flow from rich to poor countries? *Econometrica*, 84(4):1477–1521.

Comin, D. and Nanda, R. (2019). Financial development and technology diffusion. *IMF Economic Review*, 67(2):395–419.

Conway, P., De Rosa, D., Nicoletti, G., and Steiner, F. (2007). Product market regulation and productivity convergence. *OECD Economic Studies*, 2006(2):39–76.

Corrado, C., Haskel, J., and Jona-Lasinio, C. (2017). Knowledge Spillovers, ICT and Productivity Growth. *Oxford Bulletin of Economics and Statistics*, 79(4):592–618.

Crafts, N. (2006). Regulation and productivity performance. *Oxford Review of Economic Policy*, 22(2):186–202.

De la Fuente, A. and Marín, J. M. (1996). Innovation, Bank Monitoring, and Endogenous Financial Development. *Journal of Monetary Economics*, 38(2):269–301.

Denison, E. F. (1967). *Why growth rates differ: Postwar experience in nine Western countries*. Washington, DC: Brookings Institution.

Eaton, J. and Kortum, S. (1999). International Technology Diffusion: Theory and Measurement. *International Economic Review*, 40(3):537–70.

Eberhardt, M., Helmers, C., and Strauss, H. (2013). Do Spillovers Matter When Estimating Private Returns to R&D? *The Review of Economics and Statistics*, 95(2):436–448.

Ertur, C. and Musolesi, A. (2017). Weak and strong cross-sectional dependence: A panel data analysis of international technology diffusion. *Journal of Applied Econometrics*, 32(3):477–503.

Everaert, G., Heylen, F., and Schoonackers, R. (2014). Fiscal policy and tfp in the oecd: measuring direct and indirect effects. *Empirical Economics*, 49(2):605–640.

Fagerberg, J., Srholec, M., and Verspagen, B. (2010). Innovation and economic development. In *Handbook of the Economics of Innovation*, volume 2, pages 833–872. Elsevier.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, 105(10):3150–82.

Fracasso, A. and Marzetti, G. V. (2015). International trade and R&D spillovers. *Journal of International Economics*, 96(1):138–149.

Frühwirth-Schnatter, S. and Wagner, H. (2010). Stochastic Model Specification Search for Gaussian and Partial Non-Gaussian State Space Models. *Journal of Econometrics*, 154:85–100.

Fuglie, K. (2017). R&D Capital, R&D Spillovers, and Productivity Growth in World Agriculture. *Applied Economic Perspectives and Policy*, 40(3):421–444.

Griffith, R., Redding, S., and Van Reenen, J. (2003). R&D and absorptive capacity: theory and empirical evidence. *The Scandinavian Journal of Economics*, 105(1):99–118.

Griffith, R., Redding, S., and Van Reenen, J. (2004). Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries. *The Review of Economics and Statistics*, 86(4):883–895.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, 10(1):92–116.

Grossman, G. M. and Helpman, E. (1991). *Innovation and Growth in the Global Economy*. Cambridge, Mass: MIT Press.

Howitt, P. (2000). Endogenous Growth and Cross-Country Income Differences. *American Economic Review*, 90(4):829–846.

Islam, N. (1995). Growth empirics: a panel data approach. *The Quarterly Journal of Economics*, 110(4):1127–1170.

Keller, W. (1996). Absorptive capacity: On the creation and acquisition of technology in development. *Journal of Development Economics*, 49(1):199–227.

Keller, W. (2004). International Technology Diffusion. *Journal of Economic Literature*, 42(3):752–82.

Keller, W. (2010). International Trade, Foreign Direct Investment, and Technology Spillovers. In Hall, B. H. and Rosenberg, N., editors, *Handbook of the Economics of Innovation*, volume 2, chapter 19, pages 793 – 829. North-Holland.

King, R. G. and Levine, R. (1993). Finance and Growth: Schumpeter Might Be Right. *The Quarterly Journal of Economics*, 108(3):717–737.

Klenow, P. J. and Rodríguez-Clare, A. (2005). Externalities and Growth. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1A, chapter 11, pages 817 – 861. Elsevier.

Kneller, R. and Stevens, P. A. (2006). Frontier technology and absorptive capacity: Evidence from OECD manufacturing industries. *Oxford Bulletin of Economics and Statistics*, 68(1):1–21.

Laeven, L., Levine, R., and Michalopoulos, S. (2015). Financial Innovation and Endogenous Growth. *Journal of Financial Intermediation*, 24(1):1–24.

Levine, R. (1997). Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature*, 35(2):688–726.

Levine, R., Loayza, N., and Beck, T. (2000). Financial intermediation and growth: Causality and causes. *Journal of Monetary Economics*, 46(1):31–77.

Lu, Y., Tao, Z., and Zhu, L. (2017). Identifying FDI spillovers. *Journal of International Economics*, 107:75–90.

Madsen, J. B. (2007). Technology Spillover through Trade and TFP Convergence: 135 Years of Evidence for the OECD Countries. *Journal of International Economics*, 72(2):464–480.

Madsen, J. B. (2008). Economic Growth, TFP Convergence and the World Export of Ideas: A Century of Evidence. *The Scandinavian Journal of Economics*, 110(1):145–167.

Madsen, J. B. and Ang, J. B. (2016). Finance-led growth in the OECD since the nineteenth century: How does financial development transmit to growth? *Review of Economics and Statistics*, 98(3):552–572.

Madsen, J. B., Islam, M., and Ang, J. B. (2010). Catching up to the technology frontier: the dichotomy between innovation and imitation. *Canadian Journal of Economics*, 43(4):1389–1411.

Mancusi, M. L. (2008). International spillovers and absorptive capacity: A cross-country cross-sector analysis based on patents and citations. *Journal of International Economics*, 76(2):155–165.

Mankiw, N. G., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2):407–37.

Nelson, R. R. and Phelps, E. S. (1966). Investment in Humans, Technological Diffusion and Economic Growth. *American Economic Review*, 56(2):67–75.

OECD (2014). Economy-wide Regulation — Online Database.

Parente, S. and Prescott, E. (1994). Barriers to Technology Adoption and Development. *Journal of Political Economy*, 102:298–321.

Parente, S. and Prescott, E. (2002). *Barriers to Riches*. MIT Press Books.

Pesaran, H. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4):967–1012.

Pottelsberghe, B. V. and Lichtenberg, F. (2001). Does Foreign Direct Investment Transfer Technology Across Borders? *The Review of Economics and Statistics*, 83(3):490–497.

Pritchett, L. (2001). Where has all the education gone? *The World Bank Economic Review*, 15(3):367–391.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development*, 22(9):1325–1343.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics*, 70(1):65–94.

Stock, J. and Watson, M. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of Macroeconomics*, volume 2, pages 415–525. El.

Tinbergen, J. (1942). Zur Theorie der langfristigen Wirtschaftsentwicklung. *Weltwirtschaftliches Archiv*, 55:511–549.

van Ark, B., O'Mahoney, M., and Timmer, M. P. (2008). The Productivity Gap between Europe and the United States: Trends and Causes. *Journal of Economic Perspectives*, 22(1):25–44.

Westerlund, J. and Urbain, J. (2015). Cross-sectional averages versus principal components. *Journal of Econometrics*, 185:372–377.