IJIE
International Journal of
Innovation and Entrepreneurship

Modelling Innovation and Growth in Panel Data

Boyu Mu [1,*] and Yue Li [2]

[1] School of Education, University of Exeter, Exeter EX4 4PY, UK
[2] School of Management, University of York, York YO10 5DD, UK; yl2997@york.ac.uk
* Corresponding author: muboyu2020@outlook.com

Submitted: 2 August 2022, accepted: 31 October 2022, published: 19 December 2022

Abstract: This paper comprehensively reviews how innovation and growth are modelled in theoretical and empirical literature. We distinguish between economic modelling (microfounded) and econometric modelling (ad hoc). The two modelling approaches are complementary to each other for their comparative advantages in causality identification and forecasting performance. Popular models of the two approaches are illustrated and compared. We also propose an eclectic approach to combine the two approaches in one analysis framework.

Keywords: innovation; growth; panel data

How to cite: Mu, B.; Li, Y. Modelling Innovation and Growth in Panel Data. *Int. J. Innov. Entrep.*, 2022, 1(1): 3; doi:10.56502/IJIE1010003.

1. Introduction

Innovation is the engine for economic growth (Romer 1990; Aghion and Howitt 1992), and a sustainable economic growth is the key to resolve many socio-economic problems such as public debt, pension deficit, and social welfare (Foreman-Peck and Zhou 2020). However, the output of innovation—knowledge and technology—has worldwide impact and permeates across borders. This paper aims to study the determinants and effects of R&D or innovation activities in the OECD countries, taking into account the dynamic inter-dependencies between the domestic and global macroeconomic environments. The current economic research in this field is divided into two disconnected strands. One adopts the theoretical **economic modelling** framework, such as Dynamic Stochastic General Equilibrium (DSGE) techniques, widely used by government policymakers to conduct the innovation policy evaluation. The other is traditional **econometric modelling** framework, such as panel vector autoregression (PVAR) techniques, widely used by empirical studies to describe the data. This paper proposes to integrate both modelling techniques in one unified empirical framework to model a panel data of innovation activities in the OECD countries. This provides robust conclusions on the role of innovation in a globalized economy and how shocks are propagated across countries. Innovation policy implications can also be drawn based on this model, which can be used to simulate the counterfactual consequences of alternative policy arrangements (Foreman-Peck and Zhou 2022).

This paper comprehensively reviews the economic and econometric modelling approaches in the existing literature. It discusses the applications in different data types like cross-sectional, time-series, and panel data contexts. The paper aims to make a methodological contribution to bridging the theoretical and empirical research on innovation by providing a model guideline.

2. Literature Review

To explore the relationships among variables of interest, one can either follow "let data speak" philosophy to *postulate* a set of relationships qualitatively consistent with some economic theory or follow "let theory speak" philosophy to *derive* a set of relationships quantitatively consistent with some economic theory. We can term the first paradigm the "econometric modelling approach" and the second the "economic modelling approach". The econometric approach encompasses both reduced-form econometric models (e.g., cross-section regression models, time-series VAR models and panel VAR models, see Canova and Ciccarelli 2013) and structural-form econometric models (e.g., cross-sectional simultaneous equation models (SEM), time-series SVAR and panel SVAR models, see Pedroni 2013). Though the structural-form models explicitly incorporate some economic theories into the econometric models, these restrictions are usually imposed piece by piece without systematic derivation. In other words, both reduced-form and structural-form econometric models are *ad hoc* in nature. By contrast, the second approach (e.g., DSGE, see Zhang and Zhou 2021) starts with well-defined economic problems, such as consumer's utility maximization and firm's profit maximization, and systematically derives all the equilibrium conditions. These conditions are *microfounded* in nature because they are coherent and consistent with the microeconomic optimization principle. In fact, the economic models can be treated as imposing a set of theoretical restrictions on econometric models, so that the structural parameters and shocks are meaningfully identified.

A fundamental disadvantage of econometric models is its lack in theoretical foundation, but this is also its advantage—relatively free from wrong theories. They are usually more useful to summarize the stylized facts from the observed data, but they have limitations in counterfactual policy analysis and forecast in the light of the famous "Lucas Critique" (Lucas 1976). Alternatively, the economic models are designed to be immune to the Critique because all parameters and shocks in economic models are definitionally identified. More importantly, the econometric models require all variables to be observable, and omission of relevant variables will induce bias in estimation. The economic models, on the other hand, do not rely on this demanding requirement—unobservable variables can be easily handled by various techniques such as Kalman filter and particle filter in estimation. Similarly, the advantage of economic models is also its disadvantage. The validity and performance of the inference based on economic models will be undermined if the economic models are misspecified.

Seeing the relative advantages and complementarity of the two approaches, in this paper we will use both *ad hoc* econometric modelling techniques—Panel VAR (PVAR), and *microfounded* economic modelling techniques—Multi-Country DSGE model. The focus in comparison is the impulse response functions derived from these models.

3. Ad Hoc Econometric Modelling

To model panel data, the simplest approach is to use fixed effects (FE) or random effects (RE) models to capture the heterogeneity in the intercept across individuals and time.

$$y_{it} = \alpha + \alpha_i + \boldsymbol{\beta}'\mathbf{x}_{it} + \varepsilon_{it}\,, \text{ where } \varepsilon_{it} \sim IID\left(0, \sigma^2\right) \tag{1}$$

- y_{it} is the innovation expenditure to be explained in country i at t;
- $\mathbf{x}_{it}$ is the state variables to explain y_{it}, probably including lags;
- α is the intercept common to all countries and years;
- α_i is the country specific fixed/random effect;
- α_t is the time specific fixed effect;
- $\boldsymbol{\beta}$ is the slope coefficient vector common to all countries and years;
- ε_{it} is the error term assumed to be uncorrelated with α_i, $\mathbf{x}_{it}$ and itself over i and t.

Two limitations exist for this naïve reduced-form econometric model: (i) the slope coefficients are assumed homogenous; (ii) the shocks are assumed homogeneous and orthogonal, which means the shock in one country has no correlation with the shock in another. In a country level panel data, both assumptions are not quite true. To fully capture the interdependencies and heterogeneities across countries over time, we need to use PVAR.

For illustration purpose, consider a simple VAR(p) without cross-sectional variation:

$$\mathbf{y}_t = \mathbf{A}(t) + \mathbf{B}(t)\mathbf{y}_{t-1} + \mathbf{C}(t)\mathbf{z}_t + \boldsymbol{\varepsilon}_t\,, \text{ where } \boldsymbol{\varepsilon}_t \sim IID\left(0, \boldsymbol{\Sigma}\right) \tag{2}$$

- $\mathbf{y}_t$ is all the N_n endogenous variables at t, e.g., GDP, inflation, exchange rate;
- $\mathbf{y}_{t-1}$ is the N_n endogenous state (predetermined) variables at t;
- $\mathbf{z}_t$ is the N_x exogenous state variables at t, e.g., oil price;
- $\mathbf{A}(t)$ contains the deterministic components: constants, seasonal dummies and trends;
- $\mathbf{B}(L)$ is a p-th degree polynomial in the lag operator L;
- $\mathbf{C}(L)$ is a q-th degree polynomial in the lag operator L;
- $\boldsymbol{\varepsilon}_t$ is the error term assumed to be uncorrelated with $\mathbf{z}_t$.

PVAR combines the spirits of model (1) and model (2). For $i \in [1, N]$ countries:

$$\mathbf{y}_t = \mathbf{A}_i(t) + \mathbf{B}_i(t)\mathbf{y}_{t-1} + \mathbf{C}_i(t)\mathbf{z}_t + \boldsymbol{\varepsilon}_{it}\,, \text{ where } \boldsymbol{\varepsilon}_{it} \sim IID\left(0, \boldsymbol{\Sigma}_i\right) \tag{3}$$

- $\mathbf{y}_{it}$ is all the N_n endogenous variables at t for country i;
- $\mathbf{y}_{t-1}$ is the $N_n N$ endogenous state variables (both domestic and foreign);
- $\mathbf{z}_t$ is the N_x exogenous state variables at t, e.g., oil price and (maybe) foreign $\mathbf{y}^*_{it}$[1];
- $\mathbf{A}_i(t)$, $\mathbf{B}_i(L)$, $\mathbf{C}_i(L)$ and $\boldsymbol{\Sigma}_{ii}$ are now all country specific.

The PVAR with p endogenous lags and q exogenous lags is denoted as PVAR(p, q). There are three differences from the simple VAR[2]: (i) **intertemporal interdependencies**: lags of all endogenous variables of all countries enter the model for country i; (ii) **intratemporal interdependencies**: $\boldsymbol{\varepsilon}_{it}$ is generally correlated across countries, i.e., $\Sigma_{ij} \equiv Cov\left[\boldsymbol{\varepsilon}_{it}, \boldsymbol{\varepsilon}_{jt}\right] \neq \mathbf{0}$ for $i \neq j$; and (iii) **cross-sectional heterogeneity**: the intercept, the slope and the variance of the error terms are country specific. As a special case, the PVAR models used by Holtz-Eakin et al. (1988) and Love and Zicchino (2006) in microeconometric literature only assume a heterogeneous intercept $\mathbf{A}_i(t)$. Also, the simple FE/RE models can be treated as a univariate special case by noting that the state variable $\mathbf{x}_{it}$ in equation (1) corresponds to the endogenous state variables $\mathbf{y}_{t-1}$ and exogenous state variables $\mathbf{z}_t$ in Equations (2) and (3).

A key difficulty of estimating the PVAR model is the "curse of dimensionality". In general, a PVAR(p) with N countries, N_n endogenous variables, N_x exogenous variables, p endogenous lags and q exogenous lags has $NN_n\left(2 + pNN_n + (q+1)\,N_x + \frac{3N_n+1}{2}\right)$ parameters[3] to be estimated. For example, a two-country-two-variable PVAR(1,0) with one exogenous variable will have 140 parameters. Other things unchanged, if $N_n = 7$ as in most closed economy DSGE models, then there will be 392 parameters. Macroeconomic data are usually quarterly, so the observations are far from enough to estimate the PVAR parameters. The only way is to impose some restrictions so that a parsimonious form can be used to reduce the parameter dimensionality. We propose two popular ways[4] to impose the restrictions: PVAR with linear factor structure and Global VAR (GVAR).

---

1 If current foreign variable $\mathbf{y}^*_{it}$ is included as an exogenous variable, we will have a Global VAR model.

2 The first two are termed as dynamic interdependencies and static interdependencies in Canova and Ciccarelli (2013), but the meanings of "dynamic" and "static" are a bit misleading and not as precise as the ones used here.

3 This is the case with constant and a linear time trend. If three seasonal dummies are included, then there will be $3NN_n$ more coefficients.

4 There are other approaches, such as Bayesian VAR (Banbura et al. 2010), dynamic factor models (Stock and Watson 2011) and FAVAR (Bernanke et al. 2005). See Canova and Ciccarelli (2013) for a comparison.

3.1. Panel VAR with Linear Factor Structure

Following Canova and Ciccarelli (2013)[5], we can rewrite Equation (3) in a simultaneous equation form by stacking the NN_n equations[6]:

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\,,\ \text{where } \boldsymbol{\varepsilon}_t \sim IID\left(0, \boldsymbol{\Sigma}\right) \tag{4}$$

- $\mathbf{y}_t \equiv [\mathbf{y}_{1t}; \mathbf{y}_{2t}; \ldots; \mathbf{y}_{Nt}]$ is stacked endogenous variables for all countries;
- $\mathbf{X}_t \equiv I_{NN_n} \otimes \left[1, t, \mathbf{y'}_{t-1}, \ldots, \mathbf{y'}_{t-p}, \mathbf{z'}_t, \mathbf{z'}_{t-1}, \ldots, \mathbf{z'}_{t-p}\right]$ is stacked state variables;
- $\boldsymbol{\beta} \equiv vec\left([\mathbf{A}, \mathbf{B}_1, \ldots \mathbf{B}_p, \mathbf{C}_0, \mathbf{C}_1, \ldots, \mathbf{C}_q]\right)$ is stacked coefficients, where the coefficient matrices are stacked vertically, e.g., $\mathbf{B}_1 \equiv [\mathbf{B}_{11}; \mathbf{B}_{21}; \ldots; \mathbf{B}_{i1}; \ldots; \mathbf{B}_{N1}]$;
- $\boldsymbol{\varepsilon}_t \equiv [\boldsymbol{\varepsilon}_{1t}; \ldots; \boldsymbol{\varepsilon}_{Nt}]$ is stacked error terms.

The length of the coefficient vector $\boldsymbol{\beta}$ is $NN_n\left(2 + pNN_n + (q+1)\,N_x\right)$, which can easily exceeds the sample size available. Noting that $\boldsymbol{\beta}$ varies with countries and with variables among others, it is therefore possible to model the variations of $\boldsymbol{\beta}$ by K parsimonious factors[7]:

$$\boldsymbol{\beta} = \mathbf{F}_1 \boldsymbol{\varphi}_1 + \mathbf{F}_2 \boldsymbol{\varphi}_2 + \ldots + \mathbf{F}_K \boldsymbol{\varphi}_K + \boldsymbol{\xi} \tag{5}$$

This is essentially a linear model of the coefficient vector $\boldsymbol{\beta}$. The "regressor" or factor $\mathbf{F}_j$ has known inputs which can be determined with respect to the corresponding element in $\boldsymbol{\beta}$, while $\boldsymbol{\varphi}_j$ captures the "factor loading" effect. For example, $\boldsymbol{\varphi}_1$ is usually the common factor across countries and variables, so it is a scalar and the corresponding $\mathbf{F}_1$ is a vector of 1s. $\boldsymbol{\varphi}_2$ could capture the factors common within countries, so its length is N and the corresponding $\mathbf{F}_2$ is an indicator matrix with columns identifying countries. $\boldsymbol{\varphi}_3$ could capture the factors common within variables, so its length is N_n and the corresponding $\mathbf{F}_3$ is an indicator matrix with rows identifying variables. One can flexibly add or omit more factors in the model to suit the scenario at hand.

Combine Equations (4) and (5), we then end up with a multivariate regression model:

$$\mathbf{y}_t = \sum\nolimits_{j=1}^{K} \widetilde{\mathbf{X}}_{jt} \boldsymbol{\varphi}_j + \widetilde{\boldsymbol{\varepsilon}}_t,\ \text{where } \widetilde{\mathbf{X}}_{jt} = \mathbf{X}_t \mathbf{F}_j \text{ and } \widetilde{\boldsymbol{\varepsilon}}_t = \boldsymbol{\varepsilon}_t + \mathbf{X}_t \boldsymbol{\xi} \tag{6}$$

This is a much more parsimonious form compared to Equation (3), since the parameters to be estimated are $\boldsymbol{\varphi}_j$, which can then be used to derive $\boldsymbol{\beta}$ via Equation (5). Classical OLS can provide the consistent estimates.

3.2. Global VAR

Following Pesaran et al. (2004) and Dees et al. (2007), the exogenous state variable $\mathbf{z}_t$ in the general PVAR Equation (3) can be divided into two parts: the common global exogenous variables, such as oil price, and the country-specific exogenous variables. In GVAR models, the country-specific exogenous variables are defined as weighted averages of corresponding foreign variables. Since the foreign variables are included in the exogenous state variables, the endogenous state variable only includes domestic endogenous variables. Equation (3) can therefore be written as a VARX(p, q) model:

$$\mathbf{y}_{it} = \mathbf{A}_i(t) + \mathbf{B}_i(L)\mathbf{y}_{t-1} + \mathbf{C1}_i(L)\mathbf{y}^*_{it} + \mathbf{C2}_i(L)\mathbf{z}_t + \boldsymbol{\varepsilon}_{it}\,,\ \text{where } \boldsymbol{\varepsilon}_{it} \sim IID\left(\mathbf{0}, \boldsymbol{\Sigma}_{ii}\right) \tag{7}$$

- $\mathbf{y}_{it}$ is all the N_n endogenous variables at t for country i;
- $\mathbf{y}_{i,t-1}$ is the N_n endogenous state variables for country i;
- $\mathbf{y}^*_{it}$ is the N_n country-specific exogenous state variables for country i;
- $\mathbf{z}_t$ is the N_x common global exogenous state variables at t, e.g., oil price;

---

[5] Note that this method is different from the FAVAR despite similar names.

[6] Note that this paper uses the Matlab conventions in defining matrix, i.e., a semicolon ";" means stacking vertically, and a comma "," means stacking horizontally. All vectors are column vectors.

[7] It is similar to principal component analysis or factor analysis, but the factors defined here are more flexible.

- $\mathbf{A}_i(t)$, $\mathbf{B}_i(L)$, $\mathbf{C}1_i(L)$, $\mathbf{C}2_i(L)$ and $\mathbf{\Sigma}_{ii}$ are all country specific.

The exogenous state variable $\mathbf{y}_{it}^*$ is defined as:

$$\mathbf{y}_{it}^* \equiv \sum_{j=1}^{N} w_{ij}\mathbf{y}_{jt}, \text{ with } w_{ii} = 0$$

The weights w_{ij} can be derived from bilateral trades or capital flows between country i and the rest of the world. Intuitively, the relative importance of country j's GDP on that of country i can be described by the trade share of country j to country i relative to the rest of the world; the relative importance of country j's investment on that of country i can be described by the capital flow share of country j to country i relative to the rest of the world.

Note that the dimensionality of the PVAR is greatly reduced, since the "intertemporal interdependencies" are now captured by a single factor $\mathbf{y}_{it}^*$. Hence, other things equal, there will be $NN_n^2\,(pN - p - q)$ less parameters to be estimated in the VARX(p, q).

To estimate GVAR, the first step is to conduct country-by-country estimation of Equation (7) or a VECMX form of it to take into account cointegration:

$$\Delta\mathbf{y}_{it} = \mathbf{A}_i(t) + \boldsymbol{\alpha}_i\left(\boldsymbol{\beta}_{\mathbf{iy}}'\mathbf{y}_{i,t-p} + \boldsymbol{\beta}_{\mathbf{iy*}}'\mathbf{y}_{i,t-p}^* + \boldsymbol{\beta}_{\mathbf{iz}}'\mathbf{z}_{i,t-p}\right) + \mathbf{C1}_i(L)\Delta\mathbf{y}_{it}^* + \mathbf{C2}_i(L)\Delta\mathbf{z}_t + \boldsymbol{\varepsilon}_{it} \tag{8}$$

The second step is then to combine the country-specific VARX into a global specification—the GVAR. Note that in this step there is no estimation involved, and it is purely algebraic derivation. Define a "selection matrix", which is an indicator matrix picking up country-specific endogenous variables from the global vector of endogenous variables: $\mathbf{y}_{it} = \mathbf{S}_i\mathbf{y}_t$. Also define a "link matrix", which consists of the weights of country-specific exogenous variables from the global vector of endogenous variables: $\mathbf{y}_{it}^* = \mathbf{W}_i\mathbf{y}_t$. Rewrite the estimated Equation (7) using the two matrices for country i:

$$\mathbf{S}_i\mathbf{y}_t = \widehat{\mathbf{A}}_i(t) + \widehat{\mathbf{B}}_i(L)\mathbf{y}_{t-1} + \widehat{\mathbf{C1}}_i(L)\mathbf{W}_i\mathbf{y}_t + \widehat{\mathbf{C2}}_i(L)\mathbf{z}_t + \boldsymbol{\varepsilon}_{it} \tag{9}$$

Rearrange Equation (9), resulting in a country-specific model:

$$\widehat{\mathbf{D}}_i\mathbf{y}_t = \widehat{\mathbf{A}}_i(t) + \widehat{\mathbf{E}}_i(L)\mathbf{y}_{t-1} + \widehat{\mathbf{C2}}_i(L)\mathbf{z}_t + \boldsymbol{\varepsilon}_{it} \tag{10}$$

Stack all country-specific models to finally get the GVAR:

$$\mathbf{y}_t = \widehat{\mathbf{F}}(t) + \widehat{\mathbf{P}}(L)\mathbf{y}_{t-1} + \widehat{\mathbf{Q}}(L)\mathbf{z}_t + \boldsymbol{\varepsilon}_t, \text{ or } \mathbf{y}_t = \mathbf{X}_t\widehat{\boldsymbol{\beta}} + \widehat{\mathbf{D}}^{-1}\boldsymbol{\varepsilon}_t \tag{11}$$

- $\mathbf{y}_t \equiv [\mathbf{y}_{1t}; \mathbf{y}_{2t}; \ldots; \mathbf{y}_{Nt}]$ is stacked endogenous variables for all countries;
- $\mathbf{X}_t \equiv I_{NN_n} \otimes [1, t, \mathbf{y'}_{t-1}, \ldots, \mathbf{y'}_{t-p}, \mathbf{z'}_t, \mathbf{z'}_{t-1}, \ldots, \mathbf{z'}_{t-p}]$ is stacked state variables;
- $\widehat{\boldsymbol{\beta}} \equiv vec\left(\left[\widehat{\mathbf{F}}, \widehat{\mathbf{P}}_1, \ldots \widehat{\mathbf{P}}_p, \widehat{\mathbf{Q}}_0, \widehat{\mathbf{Q}}_1, \ldots, \widehat{\mathbf{Q}}_q\right]\right)$ is stacked coefficients;
- $\widehat{\boldsymbol{\varepsilon}}_t$ is stacked residuals.

The estimated GVAR in Equation (11) can be used for further inferences such as impulse responses and simulation. A remarkable advantage of GVAR is that the spillover effects of shocks across countries can be explicitly described. For example, one can easily derive the effect of a productivity shock in the US on the output in the UK.

4. Microfounded Economic Modelling

There should be two components in this microfounded model. First, we need to build a relevant DSGE model with common structure to all countries in the sample; second, we need to specify some link equations to describe the technology spillover mechanism among others. The following two subsections describe the basic features of the two model components.

4.1. The DSGE Model Component

The DSGE model component can be based on Harada (2017) NK model with innovation sector, Dees et al. (2014) MCNK model, Romer (1990) model, Grossman and Helpman (1991) model, Aghion and Howitt (1992) model, or any other model with a separate R&D sector. Since our focus is innovation—a real economic activity, we could put aside the nominal/monetary aspects without losing generality[8] and employ an RBC-type DSGE model. In this scenario, endogenous variables can include output (Y_t), consumption (C_t), sector-specific employment (Ls_t) and capital (Ks_t) with $s = 1, 2$, human capital (H_t) and knowledge capital (X_t), whose variations are explained by exogenous variables including goods sector productivity ($A1_t$), R&D sector productivity ($A2_t$), capital flows and trade flows.

The multi-sector DSGE model for each country specifies the following optimization problems under rational expectations:

- Household: max $\sum_{t=0}^{\infty} \beta^t u\left(C_t, leisure_t\right)$, subject to budget and time constraints[9];
- Goods Sector: max $\Pi 1_t$, $s.\ t.\ Y_t = A1_t f1\left(K1_t, H_t L1_t, X_t\right)$;
- R&D Sector: max $\Pi 2_t$, $s.\ t.\ X_{t+1} = \left(1 - \delta_X\right) X_t + A2_t f2\left(K2_t, H_t L2_t, X_t, X_t^*\right)$.

Each optimization problem will result in a set of first order conditions describing the equilibrium relationships between the endogenous variables and exogenous variables.

4.2. The Link Equation Component

In essence, the link equations are used to associate variables across countries to determine the variables which are treated as exogenous from each specific country's perspective but endogenous from the global perspective.

Firstly, the production function in R&D sector is assumed to be dependent of both domestic and foreign knowledge capital stocks. The "foreign" concept here is the same as that in GVAR, i.e., each country is exposed to a different subset of global knowledge capital. It is reasonable to assume that this knowledge capital X_t^* is a weighted average of all its trade partners.

Moreover, note that currently the trade flows and capital flows, along with exchange rates, are treated as exogenous to keep things simple. To fully endogenize these variables, we need to redefine C_t as a composite consumption including both domestic goods (C_{Ht}) and foreign goods (C_{Ft}), and also set up a new sector—trading firms as in Monacelli (2003). In this extended case, we need to specify how the exchange rate is determined such as PPP or UIP.

After the two structural components are specified, we will have N equations for the N endogenous variables. There are well-established techniques to analyse this dynamic stochastic equation system which is nonstationary and nonlinear. First, to deal with nonstationarity, we can either apply HP filtering or (preferably) divide the nonstationary variables by a basis variable to convert them into stationary ratios or rates. Second, to deal with nonlinearity, a standard procedure is to log-linearize the stationarized nonlinear system around the steady state and use perturbation method to solve the linearized equation system. Note that the structural parameters need to be calibrated using some initial values θ_0, so that the equation system can be solved numerically. After obtaining the linearized solution of the endogenous variables in terms of the state variables, we can then estimate the best $\hat{\theta}$ following ML procedure or Bayesian procedure. Alternatively, we can also use the Indirect Inference technique as in Chen et al. (2017). Indirect inference is a simulation-based empirical test for DSGE models using reduced-form econometric regressions as auxiliary models. In our case, we can use the GVAR developed in the previous section as the auxiliary model for our purpose.

Based on the estimated value of the structural parameters, we can then draw inferences such as impulse

8 If we try to explain the relationship between innovation and economic growth, then nominal variables would have limited roles according to New Classical neutrality and super-neutrality propositions.

9 Human capital investment is part of the household choice, since we have data measuring human capital.

response functions to be compared with those in the econometric models. However, to make meaningful comparison, we need to impose proper restrictions to identify structural shocks from PVAR.

5. Conclusion

This paper develops an eclectic approach to modelling panel data on global innovation, so that we can unify two strands of empirical methods. The research expects to have two potential contributions to the literature on innovation, one descriptive and the other normative. First, by using both economic and econometric modelling approaches, we can have a more comprehensive and robust understanding of how innovation occurs and disseminates in a globalized context. It provides a certain degree of identification and freedom of specification. Second, thanks to the identification provided by the DSGE model with innovation, we can simulate consequences of alternative innovation policies, not only the long run equilibrium effects, but also the short-run transitional dynamics. It will be a useful apparatus for policymakers to obtain information on how the economy reacts to innovation and innovation policies.

Author Contributions: Conceptualization, B.M., Y.L.; methodology, B.M., Y.L.; software, B.M., Y.L.; validation, B.M., Y.L.; formal analysis, B.M., Y.L.; investigation, B.M., Y.L.; resources, B.M., Y.L.; data curation, B.M., Y.L.; writing—original draft preparation, B.M., Y.L.; writing—review and editing, B.M., Y.L.; visualization, B.M., Y.L.; supervision, B.M., Y.L.; project administration, B.M., Y.L.; funding acquisition, B.M., Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We are grateful for the constructive comments from the reviewer Dr Nikolaos Tzivanakis.

Conflicts of Interest: The authors declare no conflict of interest.

References

Aghion, Philippe, and Peter Howitt. 1992. A Model of Growth Through Creative Destruction. *Econometrica* 60 (2): 323–51. [CrossRef]

Banbura, Marta, Domenico Giannone, and Lucrezia Reichlin. 2010. Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics* 25 (1): 71–92. [CrossRef]

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz. 2005. Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach. *The Quarterly Journal of Economics* 120: 387–422.

Canova, Fabio, and Matteo Ciccarelli. 2013. *Panel Vector Autoregressive Models*. European Central Bank Working Paper Series; Frankfurt am Main: European Central Bank.

Chen, Xiaodong, Patrick Minford, Kun Tian, and Peng Zhou. 2017. Who Provides the Capital for Chinese Growth: The Public or the Private Sector? *Applied Economics* 49: 2238–52. [CrossRef]

Dees, Stephane, Filippo di Mauro, M. Hashem Pesaran, and L. Vanessa Smith. 2007. Exploring the International Linkages of the Euro Area: A Global VAR Analysis. *Journal of Applied Econometrics* 22: 1–38. [CrossRef]

Dees, Stephane, M. Hashem Pesaran, L. Vanessa Smith, and Ron P. Smith. 2014. Constructing Multi-Country Rational Expectations Models. *Oxford Bulletin of Economics and Statistics* 76: 812–40. [CrossRef]

Foreman-Peck, James, and Peng Zhou. 2020. Devolving Fiscal Policy: Migration and Tax Yield. *Regional Studies* 54: 308–17. [CrossRef]

Foreman-Peck, James, and Peng Zhou. 2022. R&D subsidies and productivity in eastern European countries. *Economic Systems* 100978. [CrossRef]

Harada, Tsutomu. 2017. Endogenous innovation under New Keynesian Dynamic Stochastic General Equilibrium Model. *Economics of Innovation and New Technology* 27: 361–76. [CrossRef]

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. 1988. Estimating Vector Autoregressions with Panel Data. *Econometrica* 56: 1371–95. [CrossRef]

Love, Inessa, and Lea Zicchino. 2006. Financial Development and Dynamic Investment Behavior: Evidence from Panel VAR. *The Quarterly Review of Economics and Finance* 46 (2): 190–S210. [CrossRef]

Lucas, Robert. 1976. Econometric Policy Evaluation: A Critique. In *Carnegie-Rochester Conference Series on Public Policy*. Edited by K. Brunner and A. Meltzer. Amsterdam: Elsevier Science Publishers B.V., pp. 257–84.

Monacelli, Tommaso. 2003. *Monetary Policy in a Low Pass-through Environment*. European Central Bank Working Paper Series; Frankfurt am Main: European Central Bank.

Pedroni, Peter. 2013. Structural Panel Vars. *Econometrics* 1 (2): 180–206. [CrossRef]

Pesaran, Hashem, Til Schuermann, and Scott M. Weiner. 2004. Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model. *Journal of Business & Economic Statistics* 22: 129–62.

Romer, Paul. 1990. Endogenous Technological Change. *Journal of Political Economy* 98 (5): S71–sS102. [CrossRef]

Stock, James, and M. Watson. 2011. Dynamic Factor Models. In *Oxford Handbook of Forecasting*. Edited by Michael P. Clements and David F. Hendry. Oxford: Oxford University Press.

Zhang, Bo, and Peng Zhou. 2021. Financial development and economic growth in a microfounded small open economy model. *North American Journal of Economics and Finance* 58: 101544. [CrossRef]