

ORIGINAL RESEARCH ARTICLE

Macro-Level Determinants of Entrepreneurship and Endogeneity Bias – A Methodological Contribution

Brahim Gaies* and Adnane Maalaoui

IPAG Lab, IPAG Business School, Paris, France

Abstract

The eclectic theory of entrepreneurship has identified several macro-determinants of national entrepreneurial activities. Taking advantage of the availability of new databases, several recent empirical studies have sought to test these determinants in multicountry studies using multivariate regression models. Due to the lack of consensus around their results, this paper posits that this empirical literature may be subject to endogeneity bias, which seriously threatens its accuracy, consistency, and reliability, as well as the effectiveness of the resulting management and policy recommendations. Consequently, we methodologically demonstrate why and how endogeneity occurs in these studies by analyzing their empirical and theoretical models. We also provide a step-by-step guide to help researchers understand how to detect and correct endogeneity using IV techniques applied to a panel data analysis of the macro-determinants of early-stage entrepreneurship in a sample of 48 countries between 2000 and 2019. A 'toolkit' of generic STATA software commands specifying the tests, methods, and assumptions performed in this analysis is included. In doing so, we aim to raise awareness of endogeneity bias among researchers and to empirically guide future studies in order to avoid its hazards. Finally, after correcting for endogeneity, our analysis identifies the protection of property rights, entrepreneurial culture, income, and economic development as the most consistent macro-determinants of early-stage entrepreneurship, providing important policy and business insights.

Keywords: *System GMM; Entrepreneurship; Endogeneity; Panel data analysis*

Handling Editor: Vittoria Scalera; Received: 6 September 2020; Accepted: 1 June 2021; Published: 15 September 2022

While the neoclassical literature economic development is often linked to labor and capital (Solow, 1956), more recent studies have shown the role of entrepreneurship in promoting it by contributing to increased competition and innovation (Sarasvathy, 2014; Tomio & Amal, 2015). Thus, in order to achieve their economic and social objectives, including increasing GDP growth, reducing unemployment, narrowing income gaps, enhancing innovation, etc., policy-makers have become increasingly interested in fostering entrepreneurial activities in their respective countries, as they are convinced that a high level of entrepreneurial activities has a significant and positive impact on welfare performance at the macro level (Demil et al., 2018). Therefore, the question that arises here is how to promote entrepreneurship at the country level, or, more specifically, what factors can encourage or discourage entrepreneurial activities. The eclectic theory of entrepreneurship developed by Verheul et al. (2002) has addressed

this issue by identifying a number of determinants of entrepreneurship at the macro level. When Verheul et al. (2002) established the cornerstones of the eclectic theory, they demonstrated that the determinants of entrepreneurship at the macro level can be classified into two categories: the demand-side factors of entrepreneurship and the supply-side factors of entrepreneurship. The demand side represents opportunities for entrepreneurship, including technological development, the level of globalization, and economic development. On the other hand, the supply side represents the characteristics of the population in terms of resources and abilities that could potentially condition the number of entrepreneurs at the national level. Verheul et al. (2002) pointed out that cultural and institutional environments influence the supply side of entrepreneurship, and therefore, both Freytag and Thurik (2007) and Thai and Turkina (2014) included cultural and institutional factors in the macro-level determinants of entrepreneurship.

*Corresponding author: Brahim Gaies, Email: gaies_brahim@yahoo.fr

Despite the clear structure and depiction of this theoretical framework explaining entrepreneurship at the macro-level, no consensus has been reached on its empirical validity. Indeed, with the availability of macro data, such as the Global Entrepreneurship Monitor (GEM) and the development of regression techniques (Chowdhury et al., 2019), a dense and voluminous empirical literature on the macro-level determinants of entrepreneurship has emerged (see Table 1). However, many of these empirical studies not only are inconclusive and contradictory with respect to the sign – which is sometimes positive, sometimes negative – of the relationship between each macro-level determinant of entrepreneurship and the national level of entrepreneurship but have also failed to identify a significant link between them. The ambiguities between the level of economic and technological developments, unemployment, and aggregate entrepreneurship are some of the main examples that illustrate the conflicting relationship between entrepreneurship and its theoretical macro-determinants, as highlighted in numerous literature reviews (e.g., Arin et al. 2015; Dvouletý, 2017, 2018; Rusu & Roman, 2017). Recently, a few studies have explored the reasons behind these conflicting relationships, most highlighted several methodological shortcomings (e.g., Arin et al. 2015; Dvouletý, 2018). For example, when it comes to methodological choices, Bjørnskov and Foss (2016) pointed out that empirical studies on entrepreneurship can be biased by problems of causality, unobserved heterogeneity, and omitted variables.

To extend this literature, we show that the lack of consensus on the macro-determinants of entrepreneurship can be attributed, in large part, to the endogeneity bias that could tarnish previous empirical analyses. Endogeneity is a problem that could arise in regression analyses based on cross section, time series, or panel data when one (or all) of the explanatory variable(s) is (are) strongly correlated with the error term. If left unaddressed, the result is biased, and inconsistent inferences, theoretical misinterpretations, and misleading management and policy recommendations (Wooldridge, 2016) can occur. Endogeneity bias is mostly recognized in economics and finance (Wintoki et al., 2012), but it is beginning to be taken more seriously in studies in marketing, operations management, corporate governance, and on the resource-based view since the emergence of certain methodological papers that specifically deal with it: Zafarian et al. (2017), Ketokivi and McIntosh (2017), Ullah et al. (2018), Sande and Ghosh (2018), *inter alia*. To the best of our knowledge, to date there have been no methodological studies that have addressed this issue in the entrepreneurship literature and more specifically in empirical research on the determinants of entrepreneurship at the macro level. This paper contributes to filling this gap. It establishes a comprehensive framework for dealing with endogeneity bias that is adapted to empirical studies

Table 1. Previous studies on macro-level determinants of entrepreneurship

Panel data studies	Correcting endogeneity: Yes /No – Methods
Grilo and Thurik (2004)	No – Methods: Ordinary least squares (OLS)
Wennekers et al. (2005)	No – Methods: OLS
Sobel et al. (2007)	No – Methods: OLS
Wennekers et al. (2007)	No – Methods: OLS
Freytag and Thurik (2007)	No – Methods: OLS
Bjørnskov and Foss (2008)	No – Methods: OLS
Boissin et al. (2009)	No – Methods: OLS
Edelman and Yli-Renko (2010)	No – Methods: Structural equation modeling (SEM)
Klapper and Love (2010)	No – Methods: Generalized least squares (GLS)
Powell and Rodet (2012)	No – Methods: OLS
Valdez and Richardson (2013)	No – Methods: OLS
Thai and Turkina (2014)	No – Methods: Partial least squares (PLS)
Cullen et al. (2014)	No – Methods: GLS
Arin et al. (2015)	No – Methods: Bayesian model averaging (BMA) and OLS
Hall et al. (2016)	No – Methods: OLS
Roman et al. (2017)	No – Methods: Fixed effects model (FE)
Rusu and Roman (2017)	No – Methods: FE
Dvouletý (2017)	No – Methods: FE
Dvouletý (2018)	No – Methods: OLS
Ciešlik et al. (2018)	No – Methods: OLS
Chowdhury et al. (2019)	Yes – Methods: FE, random effects model (RE) and two-stage least squares (2SLS)
Hechavarría and Ingram (2019)	Yes – Methods: System generalized method of moments (SGMM)
Inekwe (2020)	Yes – Methods: FE and SGMM
Gaies et al. (2021)	Yes – Methods: FE, RE, GLS, and 2SLS

on the macro-level determinants of entrepreneurship by (1) identifying endogeneity bias and its manifestations and sources, (2) providing a methodological procedure to detect and correct it with different IV techniques, (3) illustrating this step-by-step procedure through a panel data analysis of the macro-determinants of early-stage entrepreneurship in a sample of 48 countries between 2000 and 2019, and, finally, (4) developing a ‘toolkit’ of generic STATA software commands specifying the tests, methods, and assumptions performed in this analysis.

Macro-level determinants of entrepreneurship and endogeneity: What sources?

In order to test the theoretical framework constructed by the eclectic theory regarding the determinants of

entrepreneurship at the macro level, empirical studies have generally been based on econometrics. More specifically, they have used regression techniques to pinpoint causal inferences confirming or invalidating the theoretical relationship between aggregate entrepreneurship and its potential macro-determinants. One of the first empirical studies focusing on this issue was conducted by Grilo and Thurik (2004); the authors examined the demand and supply sides of the macro-level determinants of entrepreneurship as classified by the eclectic theory (Thai & Turkina, 2014; Verheul et al., 2002). Grilo and Thurik's model includes indicators of population characteristics, economic and technological development, and globalization. The authors obtained a counter-intuitive result, indicating that funding constraints do not have a significant impact on entrepreneurship. Along the same vein, Wennekers et al. (2005) focused on the influence of income level and education structure on entrepreneurship at the macro level. They found that the national level of entrepreneurship differs across countries depending on their level of economic development, and, at the same time, the level of economic development depends on the national level of entrepreneurship. Like Grilo and Thurik (2004), the authors obtained counter-intuitive findings, demonstrating a negative impact of the level of GDP per capita on entrepreneurship. Similarly, to these pioneering empirical studies, more recent analyses (for an extensive literature review, see Arin et al., 2015; Dvouletý, 2017, 2018; Rusu & Roman, 2017) have revealed that the level of entrepreneurial activity systematically varies across countries depending on several macro-determinants. However, although the previous literature has emphasized the importance of these macro-determinants, it has often provided inconclusive and contradictory results because it has not addressed endogeneity bias, among other things (see Table 1).

Analyzing this literature, Bjørnskov and Foss (2016) conclude that the majority of the empirical studies on macro-determinants of entrepreneurship suffer from three main sources of endogeneity: simultaneous causality, omitted variables, and unobserved heterogeneity. In the following section, we discuss why and how these problems arise in this literature.

Simultaneous causality

Generally, studies on the macro-level determinants of entrepreneurship (e.g., Chowdhury et al., 2019; Cullen et al., 2014; Dvouletý, 2017; Gaies et al., 2021; Rusu & Roman, 2017; Tomio & Amal, 2015; Valdez & Richardson, 2013) applied the following model¹ to examine the macro factors that potentially impact entrepreneurial activities at the country level.

¹ Most notations and conventions are in accordance with those used by Wooldridge (2016). For more details, the reader can consult these references.

$$ENT_{it} = \alpha_0 + \alpha_1 det_{it}^1 + \alpha_2 det_{it}^2 + \alpha_3 det_{it}^3 + \alpha_4 det_{it}^p \dots + \zeta_{it} \quad (1)$$

$$\begin{pmatrix} ENT_{i1} \\ ENT_{i2} \\ \vdots \\ ENT_{iT} \end{pmatrix} = \begin{pmatrix} 1 & det_{i1}^1 & det_{i1}^2 & \dots & det_{i1}^p \\ 1 & det_{i2}^1 & det_{i2}^2 & \vdots & det_{i2}^p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & det_{iT}^1 & det_{iT}^2 & \dots & det_{iT}^p \end{pmatrix} \begin{pmatrix} \alpha^0 \\ \alpha^1 \\ \vdots \\ \alpha^p \end{pmatrix} + \begin{pmatrix} \zeta_{i1} \\ \zeta_{i2} \\ \vdots \\ \zeta_{iT} \end{pmatrix} \quad (1')$$

$$ENT = DET \alpha + \zeta \quad (1'')$$

ENT represents the vector of the dependent variable, which is the indicator of entrepreneurial activities at the country level. *DET* is the matrix of the explanatory variables with the intercept. α is the vector of the parameters, while ζ is the vector of the error term. $i = 1, 2, \dots, N$ indexes countries and $t = 1, 2, \dots, T$ indexes time periods. Finally, P indexes the number of explanatory variables.

Simultaneous causality occurs when the dependent variable explains or causes one of the independent variables, while the model assumes the opposite relationship. Taking Equation (1'') below and assuming that there is a simultaneous causality between the indicator of entrepreneurial activities at the country level (*ENT*) and the potential macro-level determinants of entrepreneurship (*DET*), we obtain:

$$ENT = DET\alpha + \zeta \quad (1'')$$

And at the same time,

$$DET = ENT\beta + \varepsilon, \quad (2)$$

where β is a vector of the parameters uncorrelated with α and ε is a vector of the error term uncorrelated with ζ .

This implies the following bias in the estimated parameters ($\hat{\alpha}$):

$$\hat{\alpha} = \frac{cov(DET, ENT)}{var(DET)} = \frac{cov(DET, DET\alpha + \zeta)}{var(DET)} = \alpha + \underbrace{\frac{cov(DET, \zeta)}{var(DET)}}_{\neq 0} \quad (3)$$

Where,

$$\frac{cov(DET, \zeta)}{var(DET)} = \frac{\beta(1-\beta)var(\zeta)}{\beta^2 var(\zeta) + var(\varepsilon)} \quad (4)$$

Finally, we can see in Equations (3) and (4) that the simultaneous causality problem leads to endogeneity bias, since we obtain:

$$cov(DET/\zeta) \neq 0 \quad (5)$$

Thus, Equation (5) indicates that one (or all) of the explanatory variables is (are) strongly correlated with the error term. In empirical studies on the macro-level determinants of entrepreneurship, the problem of simultaneous causality may arise because of the possibility of a 'chicken-or-egg causality dilemma' between the level of entrepreneurial activity and its presumed determinants, as identified by the eclectic theory, especially economic opportunities, technological development, and unemployment.

Economic development and entrepreneurship

According to Naudé (2011), economic development is mainly reflected by GDP and per capita income growth, as well as performance in the labor market (productivity), and can be defined as the lasting enhancement of a given society's material prosperity. In empirical studies on entrepreneurship, while the direct link between GDP per capita and macro-level entrepreneurship seems to be undisputed (Audretsch, 2007; Pinget et al., 2015), there are nuances in the nature and significance of this relationship depending on a country's level of GDP. The academic literature has shown that the entrepreneurship level in a given country is highly dependent on its economic growth and development at the local, regional, and national levels (Carree et al., 2007b). In this sense, there is a U-shaped link between the level of entrepreneurship and GDP per capita, implying that the entrepreneurship rate is high in low-income and high-income economies, whereas it is at its lowest level in countries with incomes ranging in between (Arin et al. 2015; Dvoutý, 2017). In addition, Dejardin (2000) argues that economic growth might be related to a change in an individual's assessment of certain professions and the associated profits. This change in perception may also affect entrepreneurship, which, first and foremost, is a key driver of economic growth.

In summary, it seems that economic development (GDP per capita growth and productivity) is the cause and consequence of entrepreneurship, and there is no consensus on the causality between these phenomena. This ambiguous relationship could be extended to financial development, economic integration, and human capital, given the strong correlation between these macro factors. In fact, according to Thai and Turkina (2014), countries with high levels of economic development are characterized by high-quality human capital and sophisticated financial systems that promote foreign trade and enable them to achieve fruitful economic integration, all of which provide good economic opportunities to their entrepreneurs.

Technological development and entrepreneurship

When it comes to the relationship between technological development and entrepreneurship, ambiguous results have

been highlighted by scholars. On the one hand, several empirical studies have determined that small-sized enterprises have a considerable impact on how innovation is developed and spread (Wennekers et al., 2002). This mechanism can be explained by the fact that technological change stimulates economic growth via productivity increases (Acs & Audretsch, 2005), which translates into additional opportunities for profit, thereby enhancing entrepreneurship. Theoretically, it was Joseph Schumpeter (1934) who first stipulated how relevant the existence of opportunities is; this is conditioned by the development of new knowledge (Kirzner, 1973) such as technological change, which is why evolutions in terms of R&D can be a source of opportunities (Casson, 1995). The creation of innovations – based on R&D activities, among others – is often favored by new businesses. This idea implies that entrepreneurs are essential for technological development, which, in turn, results in economic development that stimulates opportunities. On the other hand, several studies have proven that entrepreneurship is a favorable element for technological development (e.g., Kirchoff et al., 2007). For instance, according to Granstrand and Alänge (1995), entrepreneurs are decisive in the existence and success of innovation by recognizing or setting up and then seizing opportunities through the creation of new businesses. Given the disruptive nature of innovation that can affect an established company's competitive advantage, it is mostly nascent entrepreneurship that has a positive influence on innovation.

To sum up, it seems that entrepreneurship, innovation, and technological development constitute a virtuous cycle, which is manifested in positive bidirectional relationships. In fact, in addition to being positively influenced by the level of technological development, entrepreneurship and innovation are among the main factors that promote this development. For Galindo and Méndez (2014), this virtuous cycle is highly relevant for the creation of economic policies.

Unemployment and entrepreneurship

As with economic and technological development, the relationship between unemployment and entrepreneurship is fraught with ambiguity. In fact, both Highfield and Smiley (1987) and Evans and Leighton (1990) concluded that in the context of unemployment, individuals are more likely to set up their own business, which points to a positive relationship between unemployment and entrepreneurship. However, research has not reached a consensus on the nature of this relationship, with other findings suggesting that it may be negative (Audretsch and Fritsch, 1994). Thus, the link between unemployment and entrepreneurship has been described in two different ways, distinguishing a push- and a pull-effect (Carree et al., 2007a). The former implies that, in the context of

unemployment, individuals lack opportunities for employment, which is why they are 'pushed' toward self-employment and may set up new businesses out of necessity. Many studies have revealed this positive effect of unemployment on business creation (Blanchflower & Meyer, 1994). By contrast, the pull-effect suggests that entrepreneurship is more associated with prosperous economic conditions, meaning that, first, at low unemployment rates, individuals are more likely to turn toward entrepreneurship because they have the opportunity to resume paid employment (Muehlberger, 2007), and, second, the number of entrepreneurs tends to be low when the unemployment rate is high (Jovanovic, 1982).

In light of this, there seems to be a bidirectional relationship between unemployment and entrepreneurship, as is the case with economic and technological development and entrepreneurship. This is the main source of the simultaneous causality problem in studies on macro-level determinants of entrepreneurship, which could explain their ambiguous results.

Institutions, culture, and entrepreneurship

As is the case with the other macro-determinants of entrepreneurship, the relationship between entrepreneurial culture and aggregate entrepreneurship could be tarnished by simultaneous causality. On the one hand, it goes without saying that sound institutions, including the appropriate protection of private property rights, promote entrepreneurial activities because they reduce economic risk, uncertainty, and transaction costs (Redford, 2020). On the other hand, Henrekson and Sanandaji (2011) demonstrate that entrepreneurs can influence institutions, since they can challenge, evade, change, or accept them. Furthermore, they can create new institutions by developing innovative political activities. In addition, countries with a high culture of entrepreneurship are likely to be characterized by a similarly high level of entrepreneurial activity (Valdez & Richardson, 2013). From another perspective, the higher the level of entrepreneurial capital at the national level, the greater the national legitimacy of the entrepreneur, allowing for a strong culture of support for entrepreneurship (Kibler et al., 2014). More generally, a social environment that values entrepreneurship also has a positive impact on a country's level of nascent entrepreneurship (Mai & Gan, 2007). Additionally, regarding entrepreneurial role models, Bosma et al. (2011) suggest that four different kinds of influence can be exerted, namely, example-based learning, supportive learning, confidence, and motivation/inspiration.

Omission of variables

As its name implies, the omission of variables describes an empirical model that does not include one (or more)

important independent variable(s) to explain the phenomenon under investigation. Thus, the omitted variable(s) remain(s) in the error term and could, therefore, correlate with the independent variables, implying an endogeneity bias (Ketokivi & McIntosh, 2017; Wooldridge, 2016).

Assuming that Equation (1'') does not include one of the key independent variables, while model (6) does, we obtain:

$$ENT = DET\alpha + \zeta \quad (1'')$$

And,

$$ENT = DET\alpha + X\beta + \mu \quad (6)$$

where μ is the vector of the error term, X is the vector of the omitted variable, and β is the related vector of parameters.

We obtain,

$$\zeta = X\beta + \mu$$

Thus, if $Cov(X_{it}/DET_{it}) \neq 0$, there is an endogeneity bias, such that:

$$cov(DET/\zeta) \neq 0$$

Previous empirical studies on macro-level determinants of entrepreneurship could suffer from the omission of variables, as they do not consider past realizations of entrepreneurship to explain its current level. In the static panel data specifications commonly used by these studies, such as in Equation (1), the model does not correctly display the linkages between past and present levels of entrepreneurship and even implies that current levels are not related to past ones. Consequently, the absence of a lagged variable of a country's entrepreneurship level among the independent variables renders a specification incapable of reflecting the true reality of the entrepreneurship phenomenon and, therefore, causes an endogeneity bias, since several theoretical studies have proven the dynamic nature of entrepreneurship. Indeed, Holcombe (1998) suggests that existing entrepreneurial activities might inspire other entrepreneurs to create and seize opportunities for new businesses. Equally, Minniti (2005) has suggested that entrepreneurship enhances itself, as entrepreneurs can function as role models, hence motivating individuals to engage in entrepreneurial activity (Arenius & Minniti, 2005). In addition, in micro-economic theories, the national level of entrepreneurship is associated with competition, potentially hindering individuals from engaging in entrepreneurial activities. This is in line with Bain (1956), among others, and their theory of barriers to entry, meaning that a great number of entrepreneurs could be perceived as a threat by individuals considering entrepreneurship.

In sum – be it the approach suggesting that present entrepreneurship favors future entrepreneurship (Arenius & Minniti, 2005; Holcombe, 1998) or the one arguing that, due to competition,

present entrepreneurship may dissuade individuals from engaging in entrepreneurial activities (Bain, 1956) – either way, it seems that present entrepreneurship has an impact – be it positive or negative – on future entrepreneurship, which is not captured by the common specification used by empirical studies on macro-level determinants of entrepreneurship. This variable omission could be a source of endogeneity bias in these studies. By adding the past realizations of aggregate entrepreneurship into Equation (1’), the dynamic panel model that captures the determinants of entrepreneurship at the macro level can be expressed as follows:

$$ENT = ENT_{-1}\pi + DET\alpha + \zeta \quad (9)$$

where ENT_{-1} represents the vector of the past realizations of the dependent variable and π is the related vector of the parameters.

However, even if the dynamic panel model reduces the problem of the omission of variables, it can generate another endogeneity problem: so-called ‘dynamic endogeneity’ (Ullah et al., 2018; Wintoki, 2012). Past realizations of aggregate entrepreneurship can, thus, be correlated with the error term, such that:

$$cov(ENT_{-1}/\zeta) \neq 0 \quad (10)$$

Unobserved heterogeneity

According to Cumming and Li (2013), Bjørnskov and Foss (2016), and Chowdhury et al. (2019), previous studies on macro-determinants of entrepreneurship based on the Ordinary Least Squares (OLS) method applied to time series or cross-sectional data are flawed because they failed to account for individual time-invariant effects such as geography. This could imply the existence of a problem related to the omission of variables. Although this unobserved heterogeneity can be captured by the fixed effects (FEs) model (if it is a fixed parameter such as geography) and the random effects (REs) model (if it is a random variable such as a country’s type of business management), endogeneity bias can occur if the unobserved heterogeneity is correlated with the error term, such that:

$$cov(DET/i) \neq 0 \text{ where } \zeta = i + \mu \text{ and } i \text{ is the individual time-invariant effect} \quad (11)$$

Macro-level determinants of entrepreneurship and endogeneity: The remedy of instrumental variables

Given that empirical studies on the macro-determinants of entrepreneurship are based on observational data and frequently suffer from simultaneous causality, dynamic endogeneity, omitted variables, and unobserved heterogeneity, instrumental variable (IV) techniques are the most

appropriate method to address the endogeneity that might affect them (Bjørnskov & Foss, 2016), as are most empirical studies in economics and management (Ketokivi & McIntosh, 2017; Sande & Ghosh, 2018; Ullah et al., 2018; Zaefarian et al., 2017). To illustrate how the IV technique operates, we assume that there is an endogeneity bias in Equation (1’), which can be expressed as follows:

$$ENT = DET\alpha + \zeta \quad (1’)$$

And,

$$cov(DET/\zeta) \neq 0 \quad (5)$$

As shown in Equation (5), assuming that one (or more) of the determinant variables is correlated with the error term, to cope with this bias, an alternative (matrix/vector of) variable(s) Ω called ‘instrument(s)’ should replace the endogenous variable(s). It should verify two conditions. On the one hand, it should be strongly and significantly correlated with the endogenous variable, such that:

$$cov(\Omega/DET) \neq 0 \quad (12)$$

On the other hand, the instruments should not be correlated with the error term, such that:

$$cov(\Omega/\zeta) = 0 \quad (13)$$

Instrumental variable techniques for static models

According to Ketokivi and McIntosh (2017), two-stage least squares (2SLS) regression analysis is the most widely used IV technique in management science to address endogeneity bias. The 2SLS process involves two steps:

- Regress endogenous explanatory variables on instruments, using OLS:

$$DET = \Omega \eta + \varepsilon, \text{ where } \varepsilon \text{ is the vector of the error term} \quad (14)$$

Then,

- Insert the estimated endogenous explanatory variables \widehat{DET} – obtained from Equation (14) – into Equation (1’) and then estimate the latter by the OLS method:

$$ENT = \widehat{DET}\beta + \zeta \quad (15)$$

where β is the (new) vector of parameters.

Hence, we obtain:

$$cov(\widehat{DET}/\zeta) = 0 \quad (16)$$

IV techniques can make use of several estimators, including limited-information maximum likelihood (LIML) and the generalized method of moments (GMM). Such estimators follow the same steps as the 2SLS method (Ketokivi & McIntosh, 2017; Sande & Ghosh, 2018; Ullah et al., 2018; Zaefarian et al., 2017). As a commonly adopted approach to tackle endogeneity issues, the use of IVs bears a major challenge. In fact, not only is identifying the most relevant instrument quite demanding, but also, on the contrary, if the chosen instrument turns out to be a 'bad' one, the econometric model is likely to be less effective. In this case, the coefficients and resulting interpretations may be inconsistent. In other words, choosing an inappropriate instrument may be even more damaging to the estimates. The IV technique must, therefore, be applied with caution, and its use is not without difficulty. In fact, when determining the most appropriate instrument, it is crucial to identify variables that are highly correlated with the variable that is assumed to be endogenous, without being correlated with the error term. Overall, the IV technique must be applied under two main assumptions/conditions:

- There is a strong and statistically significant correlation between the instruments used in Equation (12) and the endogenous explanatory variables instrumented by these instruments.
- The instruments should not be correlated with the error term as expressed in Equation (13).

Instrumental variable techniques for dynamic models

Since the 2SLS, LIML, and GMM techniques are appropriate for static models (Equation 1), they are unable to deal with dynamic endogeneity due to the omission of variables. To overcome this limitation, Arellano and Bond (1991), Arellano and Bover (1995), Blundell and Bond (1998) developed two dynamic panel estimators based on the GMM and including the past values of the dependent variable within the explanatory variable to address the problem of omitted variables, as illustrated in Equation (9). The first estimator (difference GMM) is generated by transforming the dynamic model (Equation (9)) in first differences and then estimating it by the GMM.

$$\Delta ENT = \Delta ENT_{-1} \pi + \Delta DET \alpha + \Delta \zeta \quad (9')$$

This eliminates the individual time-invariant effect that could introduce the problem of unobserved heterogeneity, which is one of the main sources of endogeneity, as mentioned previously.

Since $\Delta \zeta = \Delta i + \Delta \mu$, where i is the individual time-invariant effect, $\Delta i = 0$, and $\Delta \mu$ is an *i.i.d.* variable. (11')

Then, the lagged levels of the explanatory variables are used as instruments for the endogenous variables expressed in first differences under the assumption of no serial correlation. However, as well evidenced by Blundell and Bond (1998), the difference GMM estimator could lack precision, and, even worse, it could generate biased results in finite samples when the series are persistent over time and when the time dimension is not very short (T) compared to the individual dimension (N). In this case, the lagged levels of the explanatory variables are poor quality instruments, as they are often weakly correlated with the first-difference instrumented variables due to the persistence of the series. This is clearly a concern with country panel data, since most country-level economic indicators, such as human capital, income level, financial development, innovation, and physical capital, are characterized by strong persistence over time, even after controlling for time trends. Moreover, cross-country macro panel data include relatively small individual and time dimensions compared to meso and micro panel data, which usually leads to a higher country variance than the variance of transitory shocks. For this reason, the first use of the GMM difference was in industrial labor market studies with hundreds and thousands of individual units (Staiger & Stock, 1997).

To improve the difference of the GMM estimator by solving the problem of weak instruments, Arellano and Bover (1995) and then Blundell and Bond (1998) developed a new method. These studies combined the estimation of the equation in levels (the original model) and in first-differences using lagged and level instruments for the differenced equation (Equation (9')) and different instruments for the equation in levels (Equation (9)). In other words, where the lagged levels of the explanatory variables are used as instruments for the endogenous variables expressed in first-differences, lagged differences are used as instruments for the endogenous variables in levels under the assumption that there is no correlation between past values of the instrumented variables and the current error term in levels, which includes the individual time-invariant effect. Hence, the authors transform the difference GMM estimator into a system GMM estimator, which can be expressed as follows:

$$\begin{cases} \Delta ENT = \Delta ENT_{-1} \pi + \Delta DET \alpha' + \Delta \zeta \\ ENT = \alpha_0 + ENT_{-1} \pi + DET \alpha' + \zeta \end{cases} \quad (17)$$

Using Monte Carlo experiments, Blundell and Bond (1998) ascertain that the system GMM (SGMM) estimator outperforms the difference GMM estimator in terms of both precision and bias in finite samples ($N > T$). This makes it more suitable for cross-country macro panel data characterized by relatively small individual and time dimensions and a high risk of series persistence. As a result, the SGMM estimator has been recommended by several econometric studies (e.g., Roodman,

2018; Windmeijer et al. 2019) and methodological analyses of endogeneity bias (Ketokivi & McIntosh, 2017; Sande & Ghosh, 2018; Ullah et al., 2018; Zaefarian et al. 2017). They are also used in several areas of economics, finance, and management, such as corporate finance and governance (e.g., Granado-Peiró & López-Gracia 2017), international macro-finance (e.g., Gaies & Nabi, 2019, 2021; Law et al., 2018), and energy modeling studies (e.g., Berk et al., 2020; Gaies et al., 2019), but remain underexploited in the literature on macro-determinants of entrepreneurship (see Table 1). Along this vein, Roodman (2009, p. 156) noted that '[...] difference GMM, often suffers from weak instrumentation. The favored alternative, system GMM, works only under arguably special circumstances'. Thus, the SGMM should be carried out under two main conditions, referred to by Blundell and Bond (1998) as 'moment conditions and restrictions on the initial conditions process'. First, the error term must not be serially correlated at the second order in the differenced equation, indicating the absence of first-order serial correlation in level. Second, the instruments must be uncorrelated with the error term. Arellano and Bond's (1991) serial correlation test (the AR2 test) is used to test the first condition. More precisely, it confronts the null hypothesis that there is no second-order autocorrelation in the error term against the alternative hypothesis of a second-order autocorrelation. The second condition is tested by Hansen's (1982) test of overidentifying restrictions. The null hypothesis of the test corresponds to the nonexistence of the correlation of the instruments with the error term against the hypothesis of correlation.

Beyond these conditions tested by the AR2 test and the Hansen test, there is another concern that could threaten the consistency of the GMM estimators – albeit more the difference GMM than the SGMM – which is the problem of 'too many instruments', as noted by Roodman (2009). The issue here is that the number of instruments increases quadratically with respect to the time dimension (T), thus leading to the overfitting of instrumented variables. Accordingly, the instruments fail to expunge the endogenous components of the instrumented variables, resulting in biased estimations. The problem of too many instruments could also weaken the Hansen test, leading to false-positive results, as instruments proliferate which superficially decreases their correlation with the error term. Consequently, Bowsher (2002) and then Roodman (2009) developed two techniques to deal with the problem surrounding an excess of instruments by reducing the dimensionality of the IV matrix. The first solution is to consider only a limited number of lags and not all available lags for the instruments, making them linear rather than quadratic in T. The weakness of this solution is that it neglects the information included in the dropped instruments. The second solution retains more information, since no instrument is dropped, but the instruments are collapsed in order to allocate one

instrument for each endogenous variable and lag, instead of generating one instrument for each endogenous variable, each time period, and each lag available for that period. This also makes the instruments linear rather than quadratic in T. A final issue related to the difference and system GMM estimators is the presence of panel cross-section heteroskedasticity that could reduce their efficiency, meaning that they are not biased but do not provide the most convergent estimates. To overcome this problem, in a two-step estimation, Windmeijer (2005) generates a robust standard covariance matrix corrected for panel cross-section heteroskedasticity and autocorrelation using optimal weighting matrices.

Macro-level determinants of entrepreneurship and endogeneity: A panel data analysis

To illustrate the perils of endogeneity bias, we provide the reader with a practical step-by-step procedure to detect and correct it on the basis of observational data. We study how entrepreneurial culture (CULT), protection of private property rights (INST), and other key macro-determinants of entrepreneurship influence the national level of early-stage entrepreneurship, including new businesses and nascent entrepreneurship (ENTR). In line with the eclectic theory of entrepreneurship (Verheul et al., 2002), the macro-determinants can be classified into demand and supply factors. Demand side factors include economic growth, financial development, economic integration, and technological development (GDPG, FINDEV, INTEG, and INNOV), while supply side factors are income, unemployment, and human capital (GDPC, UNEMP, and HUMCAP).

$$ENTR = f(CULT, INST, GDPG, FINDEV, INTEG, INNOV, GDPC, UNEMP, HUMCAP) \quad (18)$$

We use unbalanced panel data by matching three different sources: the GEM, World Development Indicators (WDI), and the Heritage Foundation (HF). The sample covers 48 developing and developed countries² spanning a period ranging from 2000 to 2019, facilitating the capture of year-to-year, cross-country, and intra-country variations in variables. Table 2 presents the names, definitions, and sources of the variables and illustrates their main descriptive statistics.

The entrepreneurial literature has reached a quasi-consensus – at least theoretically – on the meaningful impact of

² Argentina, Australia, Belgium, Brazil, Canada, Chile, China, Colombia, Croatia, Denmark, Ecuador, Finland, France, Germany, Greece, Guatemala, Hungary, Iceland, India, Iran, Islamic Rep., Ireland, Israel, Italy, Jamaica, Japan, Korea, Rep., Latvia, Malaysia, Mexico, Netherlands, Norway, Panama, Peru, Poland, Portugal, Romania, Russian Federation, Singapore, Slovenia, South Africa, Spain, Sweden, Switzerland, Thailand, Turkey, the United Kingdom, the United States, and Uruguay.

Table 2. Variables, descriptive statistics, and country list

Variable	Observation	Mean	Standard deviation	Definition	Source
ENTR	690	10.2118	6.3275	Total early-stage entrepreneurial activity (TEA) rate as a percentage of the population aged 18–64 years old	GEM
GDPG	960	2.9607	3.3414	GDP growth in percentage	WDI
GDPC	960	26418.25	21399.07	Real GDP per capita in US dollars	WDI
FINDEV	916	77.9875	44.3356	Domestic credit to private sector by banks as a percentage of GDP	WDI
INTEG	939	82.8890	57.5369	Sum of exports and imports as a percentage of GDP	WDI
INNOV	788	1.4208	1.0334	National research and development expenditure as a percentage of GDP	WDI
HUMCAP	912	76.6457	5.0173	Life expectancy at birth in years	WDI
UNEMP	960	7.6796	4.9750	National unemployment rate	WDI
CULT	602	62.185	13.3285	People who consider entrepreneurship as a good career choice as a percentage of the population aged 18–64 years old	GEM
INST	960	63.7962	23.0512	Property rights index '0' low; '100' high	HF

private property rights and entrepreneurial culture on aggregate entrepreneurship (Redford, 2020; Valdez & Richardson, 2013; Verheul et al., 2002), in contrast to the majority of macro-determinants of entrepreneurship for which previous research has produced largely inconclusive and contradictory results. On this basis, we expect a positive and statistically significant impact of CULT and INST on ENTR. However, as mentioned earlier, the relationship between entrepreneurial culture, property rights, and aggregate entrepreneurship could be tarnished by endogeneity due to simultaneous causality (Henrekson & Sanandaji, 2011; Kibler et al., 2014; Redford, 2020; Valdez & Richardson, 2013).

Furthermore, since certain theoretical models (e.g., Carree et al., 2007b) posit a U-shaped link between the level of entrepreneurship and the national production level, implying that the relationship between entrepreneurial activities and economic development could be non-linear, we check the linearity assumption of Equation (18) by comparing it with the following quadratic specification:

$$ENTR = f(CULT, INST, GDPC^2, FINDEV, INTEG, INNOV, GDPC^2, UNEMP, HUMCAP) \quad (18')$$

We use the Likelihood-ratio test (LR test) to compare the goodness-of-fit between the linear model (Equation (18)) and the quadratic model (Equation (18')). Table 3 (LR test – Linearity) reports the results of the LR test. It illustrates that the linear model (Equation (18)) fits the data significantly better than the quadratic model (Equation (18')) at the 1% level, since the test statistic is equal to 3.50, and the *p*-value is well above 10%.

In order to check whether the linear model is better specified than the quadratic model, we also use the variance inflation factor (VIF) test to detect any potential multicollinearity problems in both models. According to Table 3, there is a very high

Table 3. Linearity, multicollinearity, and heteroscedasticity tests

Linear model	Quadratic model
GDPC	(GDPC) ²
INST	GDPC
HUMCAP	INST
INNOV	HUMCAP
CULT	INNOV
FINDEV	GDPG
UNEMP	(GDPC) ²
GDPG	CULT
INTEG	FINDEV
	UNEMP
	INTEG
Variance inflation factor test – multicollinearity	
Mean VIF = 2.23	Mean VIF = 89.90
Breusch–Pagan test – heteroscedasticity	
chi ² (1) = 130.16	chi ² (1) = 108.90
Prob > chi ² = 0.0000	Prob > chi ² = 0.0000
Likelihood-ratio test – linearity	
LR chi ² (2)	3.50
Prob > chi ²	0.1742

correlation between the independent variables in the quadratic model (Equation (18')), as the VIF exceeds 10 on average (Mean VIF = 89.90), whereas the linear model (Equation (18)) shows a low correlation between the independent variables with a Mean VIF of 2.23. Thus, the VIF test reveals that the linear model is better specified than the quadratic model.

Finally, Table 3 reports the results of the Breusch-Pagan test, which determines the presence of the heteroscedasticity problem not only in the quadratic model but also in the linear model at the 1% level. We, therefore, generate robust/cluster

robust standard errors for our estimates for the remainder of the study, using the STATA commands 'robust or cluster ()'.

We begin the step-by-step procedure to detect and correct endogeneity bias by performing OLS regressions and examining the possible presumptions of endogeneity that may appear in our investigation. Next, we set up a process to detect endogeneity bias and its sources in our basic model (Equation 18) using the Durbin–Wu–Hausman test, and the FE and RE models. Finally, we apply four IV techniques (2SLS, GMM, LIML, and SGMM) to correct for endogeneity, allowing us to obtain consistent estimates highlighting the most relevant macro-predictors of aggregate early-stage entrepreneurship.

OLS regression and presumptions of endogeneity

Since OLS is the main method adopted in previous studies on the macro-determinants of entrepreneurship (see Table 1), we start by applying it to our basic model (Equation 18) using the STATA 'regress' command. Column 1 in Table 4 reports the results of this regression, demonstrating a positive and significant impact of GDPG, HUMCAP, and CULT on ENTR, but a negative and significant impact from GDPC, INNOV, INTEG, and UNEMP. Furthermore, there is no significant effect of FINDEV and INST. These results should be taken with caution. In fact, as explained previously, the OLS method generates inconsistent estimates if there is an endogeneity bias. This poses a serious threat to the reliability of our investigation, as most of our explanatory variables may be endogenous due to problems of simultaneous causality, unobserved heterogeneity, and omission of variables that cannot be corrected by the OLS method. If only one variable is endogenous, the results shown

in the first column of Table 4 could be misleading, inflated, and incorrect. Additionally, two other reasons for suspicion lead us to be skeptical about these results. First, Table 4 shows a modest level of adjusted *R*-squared (AdR2), which could indicate the presence of a problem of omission of variables, thus reinforcing the threat of endogeneity bias. Second, given the bidirectional correlation between private property rights and entrepreneurship found in previous studies (Kibler et al., 2014; Redford, 2020), the absence of a statistically significant impact of INST on ENTR can be considered an inconsistent result due to the existence of endogeneity bias. Ultimately, the OLS analysis allowed us to raise several presumptions about the existence of endogeneity bias. The next step is to statistically verify its existence.

Checking for endogeneity and its sources

In order to detect the existence of endogeneity bias among our explanatory variables, we follow the procedure specified by the Durbin–Wu–Hausman test, as suggested by Ullah et al. (2018). First, using the STATA command 'regress', we perform a regression of each explanatory variable on the remaining explanatory variables, for example, $CULT = f(INST, GDPG, FINDEV, INTEG, INNOV, GDPC, UNEMP, \text{ and } HUMCAP)$. Then, we calculate the residual obtained from each of these regressions by applying the STATA command 'predict, residuals'. Next, we alternatively include each residual as a 'new residual variable' in our basic model (Equation 18) and run OLS regressions using the STATA command 'regress'. Finally, we perform the Fisher test by introducing the STATA command 'test' after each OLS regression to see whether or

Table 4. Standard regressions

Estimator	OLS (1)	FE (2)	RE (3)	DFE (4)	DRE (5)
LEINTR				0.3928*** (0.1013)	0.8214*** (0.0414)
GDPG	0.2678*** (0.0856)	-0.0656 (0.0517)	-0.0133 (0.0522)	-0.0143 (0.0511)	0.0755*** (0.0363)
GDPC	-3.3030*** (1.4841)	-3.1162 (2.5376)	-2.4341** (1.0426)	-0.0898 (2.0488)	-0.7481** (0.3392)
FINDEV	-0.0114 (0.0091)	-0.0065 (0.0107)	-0.0077 (0.0077)	-0.0016 (0.0072)	-0.0058** (0.0024)
INTEG	-0.0145* (0.0082)	0.0161 (0.0193)	-0.0027 (0.0052)	0.0199 (0.0167)	-0.0013 (0.0016)
INNOV	-1.0832** (0.4664)	-1.1166 (1.0305)	-0.9215 (0.6148)	-0.9684 (0.6478)	-0.2141 (0.1660)
HUMCAP	0.2608* (0.1367)	-0.2716 (0.2443)	0.1281 (0.1366)	-0.0962 (0.1477)	0.0563* (0.0309)
UNEMP	-0.1749** (0.0751)	-0.2042** (0.0804)	-0.1769*** (0.0623)	-0.1194*** (0.0545)	-0.0385** (0.0189)
CULT	0.1424*** (0.0336)	0.0527* (0.0292)	0.0721*** (0.0266)	0.0357 (0.0239)	0.0329*** (0.0124)
INST	0.0626 (0.0431)	-0.0345 (0.0285)	-0.0077 (0.0242)	-0.0011 (0.0168)	0.0181 (0.0115)
Constant	12.1888 (12.7495)	60.8380* (30.9597)	22.2662** (10.8593)	12.0885 (24.3378)	2.2234 (2.6386)
AdR ² /number of countries	0.570/48	0.211/48	0.546/48	0.348/48	0.971/48
Fischer/ χ^2	20.06	4.687	350.9	18.80	5868
Time/country-effect	No/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes
Correcting heteroscedasticity	Yes	Yes	Yes	Yes	Yes

Dependent variable: ENTR. Robust standard errors in parentheses. Time effects are controlled by adding time dummies. ****p* < 0.01, ***p* < 0.05, **p* < 0.1.

not the coefficient of the new residual variable is statistically different from zero at the conventional level of significance. If this is not the case, it implies that the explanatory variable regressed in the first step (CULT in our example) is correlated with the error term in the basic model, meaning that it is endogenous.

Table 5 reports the STATA command used to implement this procedure, and Table 6 presents its results. The results reveal that most of our explanatory variables (GDPG, GDPC, INNOV, HUMCAP, and CULT) are endogenous at the 90% significance level. Given the existence of an endogeneity bias according to the Durbin-Wu-Hausman test, it seems relevant to test more sophisticated estimators than OLS. Thus, we estimate our basic model (Equation 18) using the FE estimator (STATA command: 'xtreg, fe') and the RE estimator (STATA command: 'xtreg, re').

Unlike the OLS estimator, the FE and RE estimators capture unobserved heterogeneity resulting from country-specific

characteristics, explaining why they have been increasingly used in recent studies, as shown in Table 1. In fact, the use of these estimators could reduce the problem of omitted variables since they consider unobserved heterogeneity to explain cross-country variations in the dependent variable. However, if the unobserved heterogeneity is not exogenous – that is, correlated with the explanatory variables – endogeneity bias could occur. Columns 2 and 3 of Table 4 present the results of the FE and RE regressions, respectively. A comparison of Columns 1, 2, and 3 of Table 4 reveals that the OLS, FE, and RE regression results are strikingly different. The three endogenous explanatory variables GDPG, INNOV, and CULT are statistically significant in the OLS regression but non-significant in the FE and RE regressions. Overall, the number of significant variables decreases from seven variables in the OLS regression (Column 1) to two variables (UNEMP and CULT) in the FE regression (Column 2) and to three variables (GDPC, UNEMP, and CULT) in the RE regression (Column 3). Intuitively,

Table 5. Generic STATA commands

Linearity test	Likelihood-ratio test	Linear model: regress Y X estimates store m1 Quadratic model: regress Y X' estimates store m2 lrtest m1 m2
Multicollinearity test	Variance inflation factor test	regress Y X, robust vif
Heteroscedasticity test	Breusch-Pagan test	regress Y X hettest
Standard regressions	OLS FE RE	regress Y X, robust xtreg Y X, fe robust xtreg Y X, re robust
Detecting endogeneity (DWH test)	Step 1: OLS Step 2: predict new-variable residuals Step 3: OLS Step 4: Fisher's test for residuals	regress X_endog X_exog, robust predict X_endog, residuals regress Y X_endo_residuals X_exog, robust test X_endo_residuals
Correcting endogeneity (IV techniques and tests)	2SLS LIML GMM SGMM	ivreg2 Y X_exog (X_endo = instruments), cluster (individual variable) ivreg2 Y X_exog (X_endo = instruments), liml cluster (individual variable) ivreg2 Y X_exog (X_endo = instruments), gmm2s cluster (individual variable) xtabond2 Y L.Y X, gmm (L.Y X_endo, collapse) iv(X_exog) nomata twostep robust

Table 6. Durbin-Wu-Hausman (DWH) test

Residual	GDPG	GDPC	FINDEV	INTEG	INNOV	HUMCAP	UNEMP	CULT	INST
Fisher's test for residuals $u = 0$ (statistic)	6.13	7.97	2.43	2.07	3.71	7.62	2.34	21.93	1.94
Fisher's test for residuals $u = 0$ (p-value)	0.0171	0.0071	0.1260	0.1573	0.0604	0.0083	0.1332	0.0000	0.1700
Nature of the variable	Endogenous	Endogenous	Exogenous	Exogenous	Endogenous	Endogenous	Exogenous	Endogenous	Exogenous

dramatic changes in the statistical significance of the explanatory variables from one standard estimator to another may demonstrate the persistence of endogeneity bias. This intuition is corroborated by at least two considerations.

First, the FE and RE estimators can help address endogeneity only if the unobserved heterogeneity is exogenous (Hamilton and Nickerson, 2003). This assumes that there is no significant relationship between the macro-determinants of entrepreneurship, such as economic development, economic integration, or entrepreneurial culture, as well as unobserved country-specific characteristics, such as geography, which is a relatively unrealistic assumption.

Second, the FE and RE regressions could be tarnished by endogeneity bias due to the omission of the potential dynamic relationship between the past and current levels of new businesses and nascent entrepreneurship (Holcombe, 1998). The modest level of the AdR2 for the FE and RE regressions (Columns 2 and 3 of Table 4) could be a symptom of this bias.

To test this possibility, we re-conduct the FE and RE regressions using a dynamic specification by adding the lagged value of the dependent variable (L.ENTR) among the independent variables of our basic model:

$$ENTR = f(L.ENTR, CULT, INST, GDPG, FINDEV, INTEG, INNOV, GDPC, UNEMP, HUMCAP) \quad (19)$$

The results of the dynamic FE and RE regressions are presented in Columns 4 and 5 of Table 4; these reveal a sharp increase in the level of AdR2 and a significant positive correlation between the past and present levels of early-stage entrepreneurship at the 99% level of statistical significance. While it is essential to consider these results with caution because FE

and RE techniques cannot address the dynamic endogeneity caused by the presence of the lagged value of the dependent variable (L. ENTR) among the independent variables, they could provide another presumption of the existence of the omitted variable problem, given that the dynamic nature of entrepreneurial activities is theoretically highlighted in the literature (Holcombe, 1998).

Correcting endogeneity

In order to deal with endogeneity bias, we re-conduct our static OLS, FE, and RE regressions, taking explanatory variables lagged by 1 year. According to Bellemare et al. (2017), this might address the problem of the simultaneous causality but fails to deal with the other sources of endogeneity. As displayed in Table 7, the results of the OLS, FE, and RE regressions are considerably different, as are those of the standard regressions reported in Table 4. Six lagged explanatory variables, including the three endogenous explanatory variables GDPG, INNOV, and CULT, appear statistically significant at conventional levels in the OLS regression (Column 1), evidencing a positive effect of economic growth, human capital, and entrepreneurial culture on nascent entrepreneurship, while income level, unemployment, economic integration, and technological development act as barriers to entry for new ventures. The RE regression (Column 3) confirms these results only for GDPC, CULT, INNOV, and UNEMP, while the FE regression (Column 2) reveals only one significant variable, namely, UNEMP.

Next, we performed three IV techniques – 2SLS, GMM, and LIML – with a static specification of our basic model (Equation 18) using the STATA command ‘ivreg 2’. For the dynamic specification (Equation 19), we apply the SGMM estimator using

Table 7. Regressions with lagged dependent variables

Estimator	OLS (1)	FE (2)	RE (3)
L.GDPG	0.2852*** (0.0795)	0.0213 (0.0590)	0.0670 (0.0494)
L.GDPC	-3.3881** (1.4809)	-0.7635 (4.8916)	-2.1393** (1.0546)
L.FINDEV	-0.0113 (0.0095)	-0.0053 (0.0093)	-0.0053 (0.0066)
L.CULT	0.1298*** (0.0313)	0.0160 (0.0228)	0.0413** (0.0198)
L.INST	0.0645 (0.0454)	-0.0425 (0.0299)	-0.0125 (0.0277)
L.INTEG	-0.0151* (0.0087)	0.0009 (0.0235)	-0.0088 (0.0083)
L.INNOV	-1.1209** (0.5344)	-1.4496 (1.1854)	-1.3469** (0.5906)
L.HUMCAP	0.2582* (0.1373)	-0.1192 (0.2466)	0.1834 (0.1672)
L.UNEMP	-0.1907*** (0.0704)	-0.1978* (0.1060)	-0.2121*** (0.0584)
Constant	14.2873 (12.6319)	30.4763 (50.2977)	18.5498 (11.3591)
AdR ² /number of countries	0.585/48	0.225/48	0.557/48
Fischer/ χ^2	9.553	8.183	234.9
Time/country-effect	No/Yes	Yes/Yes	Yes/Yes
Correcting heteroscedasticity	Yes	Yes	Yes

Dependent variable: ENTR. Robust standard errors in parentheses. Time effects are controlled by adding time dummies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

the STATA command 'xtabond2' as it is an IV technique that can control for dynamic endogeneity, the omission of variables, and simultaneous causality. We decided not to apply the difference GMM estimator because of its shortcomings in terms of precision and consistency in finite samples (Blundell & Bond, 1998; Staiger & Stock, 1997).

The results of the 2SLS, GMM, LIML, and SGMM regressions are reported in Table 8. For most of the coefficients of the explanatory variables, these regressions provide more similar signs and significances than the results of the standard regressions (OLS, FE, and RE) presented in Table 4, in addition to the results of the regressions with lagged dependent variables (OLS, FE, and RE) reported in Table 7. All results obtained from the 2SLS, GMM, LIML, and SGMM regressions confirm that a high level of economic development (GDPG) favors early-stage entrepreneurship by offering new opportunities, while the higher the level of income (GDPC), the higher the wages and the less attracted the population is to entrepreneurship, as explained by Wennekers et al. (2010). The same results are found with the OLS regressions but not with FE, whereas the RE regression showed the same effect for GDPC but not for GDPG. Moreover, all results from the 2SLS, GMM, LIML, and SGMM regressions confirm that CULT and INST have a positive and statistically significant impact on ENTR, in line with

previous studies (Redford, 2020; Valdez & Richardson, 2013; Verheul et al., 2002). This is also found with OLS and RE regressions for CULT but not for INST, while CULT and INST are not statistically significant in the FE regressions. Furthermore, FINDEV and INNOV have low significance (10% level of risk tolerance) in the 2SLS and GMM regressions and non-significance in the LIML and SGMM regressions, which is consistent with the OLS regression results but not with the FE and RE results.

Overall, according to the 2SLS, GMM, LIML, and SGMM regressions, after controlling for endogeneity, CULT, INST, GDPG, and GDPC appear to be the most consistent macro-determinants of early-stage entrepreneurship, with a positive impact of GDPG, CULT, and INST and a negative impact of GDPC on ENTR. This result is more congruent with the OLS regressions (regarding CULT, GDPG, and GDPC) than with the FE and RE regressions. Consequently, it indicates that OLS models can be better – or at least the 'lesser evil' – than FE and RE models when the risk of unobserved endogeneity exists, which is generally the case in studies on the macro-determinants of entrepreneurship because the national level thereof is often determined by cultural and geographical characteristics that are difficult to quantify and include in multivariate regression models without triggering an endogeneity bias.

Table 8. Instrumental variable regressions

Estimator	2SLS (1)	GMM (2)	LIML (3)	SGMM (4)
LENTRE				0.2537*** (0.0832)
GDPG	0.3909* (0.2106)	0.3822** (0.1722)	0.3843* (0.2156)	0.0764*** (0.0184)
GDPC	-2.5016** (1.2344)	-2.5717** (1.1636)	-2.5114** (1.2448)	-3.1047*** (0.5655)
FINDEV	-0.0280* (0.0155)	-0.0275* (0.0147)	-0.0284* (0.0158)	-0.0003 (0.0055)
CULT	0.5077*** (0.1780)	0.5050*** (0.1624)	0.5158*** (0.1847)	0.0638*** (0.0151)
INST	0.0646* (0.0355)	0.0675** (0.0341)	0.0648* (0.0356)	0.0735*** (0.0176)
INTEG	-0.0012 (0.0136)	-0.0001 (0.0133)	-0.0009 (0.0139)	-0.0081 (0.0057)
INNOV	1.2687 (1.6213)	1.2413 (1.4541)	1.3295 (1.6770)	-1.0358* (0.3124)
HUMCAP	0.2409* (0.1455)	0.2426* (0.1413)	0.2411 (0.1471)	0.1114 (0.0751)
UNEMP	-0.1008 (0.1170)	-0.0959 (0.1014)	-0.1003 (0.1190)	-0.1345 (0.0344)
Constant	-615.1196*** (165.5548)	-615.1144*** (162.4785)	-614.9433*** (167.2151)	-389.1789*** (79.9855)
<i>Instrument tests</i>				
LM-Anderson test (P-value)	0.05	0.05	0.05	
Sargan-Hansen test (P-value)	0.860	0.822	0.864	0.793
Number of countries	48	48	48	48
AR2 P-value				0.678
Number of instruments	14	14	14	29
AdR ²	0.742	0.744	0.737	
Fisher	5.488	5.731	5.460	
Time/country-effect	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes
Correcting heteroscedasticity	Yes	Yes	Yes	Yes

Dependent variable: ENTR. Robust standard errors in parentheses. Time effects are controlled by adding time dummies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

In addition, the positive and significant impact of the past level of early-stage entrepreneurship (L.ENTR) on its present level is confirmed by the SGMM regression at the 99% level, which is theoretically corroborated by Holcombe (1998) and entrepreneurial capital theory (Kibler et al., 2014), showing the importance of the issue of dynamic endogeneity in studies on the macro-determinants of entrepreneurship.

Furthermore, for the SGMM regression, we use the past values of the lagged endogenous variable as instruments for the first differences equation and the first-differenced values of the lagged endogenous variable with a time dummy to control for time trends, as instruments for the levels equation. In doing so, we follow the initial specification proposed by Blundell and Bond (1998) in which they developed the SGMM estimator, such that the lagged levels of the explanatory variables are the instruments of the endogenous variables expressed in first-differences, and the lagged differences are the instruments of the endogenous variables in levels. The use of a time dummy to control for time trends helps to prevent the problems associated with the persistence of the series, as pointed out by Staiger and Stock (1997). Next, to avoid the aforementioned problem of instrument proliferation, we specify one instrument per endogenous variable and per lag, rather than one instrument per time period, per endogenous variable, and per lag, using the STATA option 'collapse' with the STATA command 'xtabond2' (see Table 5). In addition, to overcome the problem of the panel cross-section heteroskedasticity, we apply the Windmeijer (2005) correction allowing the generation of a robust standard covariance matrix in the two-step estimation, as mentioned previously. We use the STATA option 'twostep' with the STATA command 'xtabond2' (see Table 5) to perform the Windmeijer correction. As shown in Table 8, the Arellano and Bond test (1991) does not reject the null hypothesis of no second-order serial correlation at the 95% level (AR2), and the Sargan-Hansen test does not reject the null hypothesis of no correlation between the instruments and the error term at the 95% level. Subsequently, these tests confirm the validity of our instruments and the results of the SGMM regression.

Conclusion

Given the potential benefits of entrepreneurship in terms of economic, technological, and social development, policy makers have an interest in identifying its determinants in order to promote it at the macro level (Acs et al. 2014). In the academic sphere, the eclectic theory of entrepreneurship responds to this challenge by providing a comprehensive theoretical framework of the macro-level determinants of entrepreneurship (e.g., Freytag & Thurik, 2007; Thai & Turkina, 2014; Verheul et al., 2002). However, despite the relevance and predominance of this theoretical framework, the empirical studies that underpin

it are characterized by contradictory findings and conclusions (Arin et al., 2015; Dvouletý, 2017, 2018; Rusu & Roman, 2017). In this paper, we introduced our analysis by arguing that the mixed empirical evidence can be attributed to methodological shortcomings (Arin et al., 2015; Dvouletý, 2018), similarly to Bjørnskov and Foss's (2016, p. 292) conclusion that the existing literature does 'not theorize many potentially relevant inter-level links and mechanisms and suffers from sample limitations, omitted variable biases, causality issues, and response heterogeneity'. In particular, we have pointed out the inability of previous studies to control for endogeneity bias – mainly caused by the omission of variables, simultaneous causality, and error-in-variables. As this statistical problem is well known in economics and finance, it is beginning to be considered more seriously in studies in marketing, operations management, corporate governance, and on the resource-based view (Sande & Ghosh, 2018; Ullah et al., 2018), but it remains unaddressed in the entrepreneurship literature, particularly in empirical studies on the macro-level determinants of entrepreneurship. A thorough analysis of the eclectic theory of entrepreneurship confronted with the empirical literature that has tested its conclusions has provided us with a cartography, proving that the main sources of endogeneity likely exist in this empirical literature.

First of all, the potential existence of a 'chicken-or-egg causality dilemma' between entrepreneurship and three of its presumed determinants identified by the eclectic theory, namely, economic development, technological development, and institutions, implies that the problem of simultaneous causality could most likely arise in previous empirical studies on the macro-level determinants of entrepreneurship.

Second, the empirical models commonly used in previous studies (static specification, observational data, OLS method) on the macro-level determinants of entrepreneurship fail to include a vital independent variable to explain entrepreneurship, namely, the lagged variable of a country's entrepreneurship level (lagged dependent variable) among the independent variables, resulting in a model that is incapable of capturing entrepreneurial dynamics. This means that, without the past realizations of entrepreneurship, the model assumes that the current levels of entrepreneurship are not related to the past levels, whereas several theoretical frameworks have shown that present entrepreneurship has an impact – be it positive or negative – on future entrepreneurship (Arenius & Minniti, 2005; Bain, 1956; Holcombe, 1998). Here, an endogeneity bias can occur due to the problem of the omission of variables.

Finally, empirical studies on the macro-level determinants of entrepreneurship could be subject to the problem of unobserved heterogeneity, which is the third source of endogeneity bias. This could occur when an unobserved fixed parameter such as geography or random variable such as a country's type of business management is correlated with the error term.

We have shown that not controlling for these three problems implies an endogeneity bias leading to incorrect and inflated results, leading to misinterpretations of estimated model parameters that could even have misleading and theoretically counter-intuitive signs. Nevertheless, previous empirical studies on the macroeconomic determinants of entrepreneurship have generally relied on standard estimators, including OLS, generalized least squares, FE, and RE that do not account for endogeneity, which seriously undermines their precision, consistency, and reliability, as well as the effectiveness of the resulting management and policy recommendations. Consequently, in starting to construct a new framework to deal with endogeneity bias, we have discussed the most appropriate and potentially usable estimators for empirical studies on the macro-level determinants of entrepreneurship that are, as shown previously, fundamentally vulnerable to the three main sources of endogeneity. Next, we study the main IV techniques, i.e., 2SLS, LIML, GMM, and SGMM, used to deal with endogeneity bias and explain how they function, the assumptions supporting their use, and the extent to which they are adapted to address this bias in the context of empirical research on the macro-level determinants of entrepreneurship. Thus, we provide a methodological procedure to detect and correct endogeneity with different IV techniques and then illustrate it with a panel data analysis of the macro-determinants of early-stage entrepreneurship in a sample of 48 countries between 2000 and 2019. In addition, we develop a 'toolkit' of generic STATA software commands specifying the tests, methods, and assumptions performed in this analysis. In doing so, we aimed to provide a comprehensive reference not only to assist empirical studies investigating the macro-level determinants of entrepreneurship to select the most suitable method to deal with the issue of endogeneity, but also to improve their quality by avoiding biased results and interpretations. Finally, by showing that economic development, income, entrepreneurial culture, and property rights are the most consistent macro-determinants of early-stage entrepreneurship after correcting for endogeneity, we aimed to provide interesting insights for policy makers, entrepreneurs, and future empirical studies on nascent entrepreneurship.

References

- Acs, Z. J. & Audretsch D. B. (2005). Entrepreneurship, innovation and technological change. *Foundations and Trends in Entrepreneurship*, 1(4), 149–195. doi: 10.1561/03000000004
- Acs, Z. J., Autio, E. & Szerb, L. (2014). National systems of entrepreneurship: Measurement issues and policy implications. *Research Policy*, 43(3), 476–494. doi: 10.1016/j.respol.2013.08.016
- Arellano, M. & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297. doi: 10.2307/2297968
- Arellano, M. & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29–51. doi: 10.1016/0304-4076(94)01642-D
- Arenius, P. & Minniti, M. (2005). Perceptual variables and nascent entrepreneurship. *Small Business Economics*, 24(3), 233–247. doi: 10.1007/s11187-005-1984-x
- Arin, K. P., Huang, V. Z., Minniti, M., Nandialath, A. M. et al. (2015). Revisiting the determinants of entrepreneurship: A Bayesian approach. *Journal of Management*, 41(2), 607–631. doi: 10.1177/0149206314558488
- Audretsch, D. B. (2007). *The entrepreneurial society*. Oxford University Press. Retrieved from <https://global.oup.com/academic/product/the-entrepreneurial-society-9780195183504?cc=fr&lang=en>
- Audretsch, D. B. & Fritsch, M. (1994). The geography of firm births in Germany. *Regional Studies*, 28(4), 359–365. doi: 10.1080/00343409412331348326
- Bain, J. S. (1956). *Barriers to new competition*. Harvard University Press.
- Bellemare, M. F., Masaki, T. & Pepinsky, T. B. (2017). Lagged explanatory variables and the estimation of causal effect. *The Journal of Politics*, 79(3), 949–963. Retrieved from <https://www.hup.harvard.edu/catalog.php?isbn=9780674188037>
- Berk, I., Kasman, A. & Kılınc, D. (2020). Towards a common renewable future: The system-GMM approach to assess the convergence in renewable energy consumption of EU countries. *Energy Economics*, 87, 103922. doi: 10.1016/j.eneco.2018.02.013
- Bjørnskov, C. & Foss, N. J. (2008). Economic freedom and entrepreneurial activity: Some cross-country evidence. *Public Choice*, 134(3–4), 307–328. doi: 10.1007/s11127-007-9229-y
- Bjørnskov, C. & Foss, N. J. (2016). Institutions, entrepreneurship, and economic growth: What do we know and what do we still need to know? *Academy of Management Perspectives*, 30(3), 292–315. doi: 10.5465/amp.2015.0135
- Blanchflower, D. G. & Meyer, B. D. (1994). A longitudinal analysis of the young self-employed in Australia and the United States. *Small Business Economics*, 6(1), 1–19. doi: 10.1007/BF01066108
- Blundell, R. & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115–143. doi: 10.1016/S0304-4076(98)00009-8
- Boissin, J. P., Chollet, B. & Emin, S. (2009). Les déterminants de l'intention de créer une entreprise chez les étudiants: un test empirique. *M@n@gement*, 12(1), 28–51. <https://management-aims.com/index.php/mgmt/article/view/4099>
- Bosma, N., Hessels, J., Schutjens, V., Praag, V. et al. (2011). *Entrepreneurship and role models*. Tinbergen Institute Discussion, Paper no. 11-061/3. Retrieved from <https://papers.tinbergen.nl/11061.pdf>
- Bowsher, C. G. (2002). On testing overidentifying restrictions in dynamic panel data models. *Economics Letters*, 77(2), 211–220. doi: 10.1016/S0165-1765(02)00130-1
- Carree, M. A., Van Stel, A., Thurik, R. & Audretsch, D. B. (2007a). Does self-employment reduce unemployment? *Entrepreneurship, Growth and Public Policy*, 1(18), 1613–8333.
- Carree, M., Van Stel, A., Thurik, R. & Wennekers, S. (2007b). The relationship between economic development and business ownership revisited. *Entrepreneurship & Regional Development*, 19(3), 281–291. doi: 10.1080/08985620701296318
- Casson, M. (1995). *Entrepreneurship and business culture*. Edward Elgar Publishing. doi: 10.1016/s0024-6301%2897%2986609-5
- Chowdhury, F., Audretsch, D. B. & Belitski, M. (2019). Institutions and entrepreneurship quality. *Entrepreneurship Theory and Practice*, 43(1), 51–81. doi: 10.1177%2F1042258718780431
- Ciešlik, J., Kaciak, E. & van Stel, A. (2018). Country-level determinants and consequences of overconfidence in the ambitious entrepreneurship

- segment. *International Small Business Journal*, 36(5), 473–499. doi: 10.1177/20266242617748201
- Cullen, J. B., Johnson, J. L. & Parboteeah, K. P. (2014). National rates of opportunity entrepreneurship activity: Insights from institutional anomie theory. *Entrepreneurship Theory and Practice*, 38(4), 775–806. doi: 10.1111/etap.12018
- Cumming, D. J. & Li, D. (2013). Public policy, entrepreneurship, and venture capital in the United States. *Journal of Corporate Finance*, 23, 345–367. doi: 10.1016/j.jcorpfin.2013.09.005
- Dejardin, M. (2000). Entrepreneurship and economic growth: An obvious conjunction. *Institute for Development Strategies*. Retrieved from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.197.3876&rep=rep1&type=pdf>
- Demil, B., Lecocq, X. & Warnier, V. (2018). 'Business model thinking', business ecosystems and platforms: The new perspective on the environment of the organization. *M@n@gement*, 21(4), 1213–1228. doi: 10.3917/mana.214.1213
- Dvouletý, O. (2017). Determinants of Nordic entrepreneurship. *Journal of Small Business and Enterprise Development*, 24(1), 12–33. doi: 10.1108/JSBED-07-2016-0104
- Dvouletý, O. (2018). How to analyse determinants of entrepreneurship and self-employment at the country level? A methodological contribution. *Journal of Business Venturing Insights*, 9, 92–99. doi: 10.1016/j.jbvi.2018.03.002
- Edelman, L. & Yli-Renko, H. (2010). The impact of environment and entrepreneurial perceptions on venture-creation efforts: Bridging the discovery and creation views of entrepreneurship. *Entrepreneurship Theory and Practice*, 34(5), 833–856. doi: 10.1111/j.1540-6520.2010.00395.x
- Evans, D. S. & Leighton, L. S. (1990). Small business formation by unemployed and employed workers. *Small Business Economics*, 2(4), 319–330. doi: 10.1007/BF00401628
- Freytag, A. & Thurik, R. (2007). Entrepreneurship and its determinants in a cross-country setting. *Journal of Evolutionary Economics*, 17(2), 117–131. doi: 10.1007/s00191-006-0044-2
- Gaies, B., Kaabia, O., Ayadi, R., Guesmi, K. et al. (2019). Financial development and energy consumption: Is the MENA region different? *Energy Policy*, 135, 111000. doi: 10.1016/j.enpol.2019.111000
- Gaies, B. & Nabi, M. S. (2019). Financial openness and growth in developing countries. *Journal of Economic Integration*, 34(3), 426–464. Retrieved from <https://www.jstor.org/stable/26756452>
- Galindo, M. Á. & Méndez, M. T. (2014). Entrepreneurship, economic growth, and innovation: Are feedback effects at work? *Journal of Business Research*, 67(5), 825–829. <https://doi.org/10.1016/j.jbusres.2013.11.052>
- Gaies, B. & Nabi, M. S. (2021). Banking crises and economic growth in developing countries: Why privileging foreign direct investment over external debt? *Bulletin of Economic Research*, 73(4), 736–761. <https://doi.org/10.1111/boer.12271>
- Gaies, B., Najar, D., Maalou, A., Kraus, S. & El Tarabishy, A. (2021). Does financial development really spur nascent entrepreneurship in Europe? A panel data analysis. *Journal of Small Business Management*, 1–48. doi: 10.1080/00472778.2021.1896722
- Granado-Peiró, N. & López-Gracia, J. (2017). Corporate governance and capital structure: A Spanish study. *European Management Review*, 14(1), 33–45. doi: 10.1111/emre.12088
- Granstrand, O. & Alänge, S. (1995). The evolution of corporate entrepreneurship in Swedish industry – Was Schumpeter wrong? *Journal of Evolutionary Economics*, 5(2), 133–156. doi: 10.1007/BF01199854
- Grilo, I. & Thurik, R. (2004). *Determinants of entrepreneurship in Europe*. ERIM Report Series Reference No. ERS-2004-106-ORG. Retrieved from <https://ssrn.com/abstract=636815>
- Hall, J. C., Lacombe, D. J. & Pokharel, S. B. (2016). Freedom and entrepreneurship: A spatial econometric approach. *Journal of Entrepreneurship and Public Policy*, 5(3), 404–411. doi: 10.1108/JEPP-12-2015-0038
- Hamilton, B. H. & Nickerson, J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization*, 1(1), 51–78. doi: 10.1177/1476127003001001218
- Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054. doi: 10.2307/1912775
- Hechavarría, D. M. & Ingram, A. E. (2019). Entrepreneurial ecosystem conditions and gendered national-level entrepreneurial activity: A 14-year panel study of GEM. *Small Business Economics*, 53(2), 431–458. doi: 10.1007/s11187-018-9994-7
- Henrekson, M. & Sanandaji, T. (2011). Entrepreneurship and the theory of taxation. *Small Business Economics*, 37(2), 167–185. doi: 10.1007/s11187-009-9242-2
- Highfield, R. & Smiley, R. (1987). New business starts and economic activity: An empirical investigation. *International Journal of Industrial Organization*, 5(1), 51–66. doi: 10.1016/0167-7187(87)90006-3
- Holcombe, R. G. (1998). Entrepreneurship and economic growth. *Quarterly Journal of Austrian Economics*, 1(2), 45–62. doi: 10.1007/s12113-998-1008-1
- Inekwe, J. N. (2020). Global financial networks and entrepreneurship. *The World Economy*, 44(5), 1261–1280. doi: 10.1111/twec.13056
- Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica: Journal of the Econometric Society*, 50(3), 649–670. doi: 10.2307/1912606
- Klapper, L. F. & Love, I. (2010). *The impact of business environment reforms on new firm registration*. Policy Research Working Paper Series 5493. The World Bank. doi: 10.1596/1813-9450-5493
- Ketokivi, M. & McIntosh, C. N. (2017). Addressing the endogeneity dilemma in operations management research: Theoretical, empirical, and pragmatic considerations. *Journal of Operations Management*, 52, 1–14. doi: 10.1016/j.jom.2017.05.001
- Kibler, E., Kautonen, T. & Fink, M. (2014). Regional social legitimacy of entrepreneurship: Implications for entrepreneurial intention and start-up behaviour. *Regional Studies*, 48(6), 995–1015. doi: 10.1080/00343404.2013.851373
- Kirchhoff, B. A., Newbert, S. L., Hasan, I. & Armington, C. (2007). The influence of university R&D expenditures on new business formations and employment growth. *Entrepreneurship Theory and Practice*, 31(4), 543–559. doi: 10.1111/j.1540-6520.2007.00187.x
- Kirzner, I. M. (1973). *Entrepreneurship and competition*. University of Chicago Press. Retrieved from [https://www.scirp.org/\(S\(43dyn45teexjx-455qlt3d2q\)\)/reference/ReferencesPapers.aspx?ReferenceID=1583665](https://www.scirp.org/(S(43dyn45teexjx-455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1583665)
- Law, S. H., Kutan, A. M. & Naseem, N. A. M. (2018). The role of institutions in finance crisis: Evidence from international data. *Journal of Comparative Economics*, 46(1), 174–191. doi: 10.1016/j.jce.2017.04.001
- Mai, Y. & Gan, Z. (2007). Entrepreneurial opportunities, capacities and entrepreneurial environments. Evidence from Chinese GEM data. *Chinese Management Studies*, 1(4), 216–224. Retrieved from <http://hdl.handle.net/10419/228732>
- Minniti, M. (2005). Entrepreneurship and network externalities. *Journal of Economic Behavior & Organization*, 57(1), 1–27. doi: 10.1016/j.jebo.2004.10.002
- Muehlberger, U. (2007). *Dependent self-employment: Workers on the border between employment and self-employment*. Springer.
- Naudé, W. (2011). *Entrepreneurship and economic development*. Palgrave Macmillan. doi: 10.1057/9780230295155
- Pinget, A., Bocquet, R. & Mothe, C. (2015). Barriers to environmental innovation in SMEs: Empirical evidence from French firms. *M@n@gement*, 18(2), 132–155. doi: 10.3917/mana.182.0132

- Powell, B. & Rodet, C. S. (2012). Praise and profits: Cultural and institutional determinants of entrepreneurship. *Journal of Private Enterprise*, 27(2), 19.
- Redford, A. (2020). Property rights, entrepreneurship, and economic development. *The Review of Austrian Economics*, 33(1), 139–161. doi: 10.1007/s11138-019-00485-6
- Roman, A., Bilan, I. & Ciumaş, C. (2018). What drives the creation of new businesses? A panel-data analysis for EU countries. *Emerging Markets Finance and Trade*, 54(3), 508–536. doi: 10.1080/1540496X.2017.1412304
- Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *The Stata journal*, 9(1), 86–136. doi: 10.1177/1536867X0900900106
- Roodman, D. (2018). *xtabond2: Stata module to extend xtabond dynamic panel data estimator*. Boston College Department of Economics.
- Rusu, V. D. & Roman, A. (2017). Entrepreneurial activity in the EU: An empirical evaluation of its determinants. *Sustainability*, 9(10), 1679. doi: 10.3390/su9101679
- Sande, J. B. & Ghosh, M. (2018). Endogeneity in survey research. *International Journal of Research in Marketing*, 35(2), 185–204. doi: 10.1016/j.ijresmar.2018.01.005
- Sarasvathy, S. (2014). The downside of entrepreneurial opportunities. *M@n@gement*, 17(4), 305–315. doi: 10.3917/mana.174.0305
- Schumpeter, J. (1934). *The theory of economic development*. Harvard University Press.
- Sobel, R. S., Clark, J. R. & Lee, D. R. (2007). Freedom, barriers to entry, entrepreneurship and economic progress. *Austrian Review of Economics*, 20(4), 221–236. doi: 10.1007/s11138-007-0023-3
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1), 65–94. doi: 10.2307/1884513
- Staiger, D. & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica: Journal of the Econometric Society*, 65(3), 557–586. doi: 10.2307/2171753
- Thai, M. T. T. & Turkina, E. (2014). Macro-level determinants of formal entrepreneurship versus informal entrepreneurship. *Journal of Business Venturing*, 29(4), 490–510. doi: 10.1016/j.jbusvent.2013.07.005
- Tomio, B. T. & Amal, M. (2015). Institutional distance and Brazilian outward foreign direct investment. *M@n@gement*, 18(1), 78–101. Retrieved from <https://management-aims.com/index.php/mgmt/article/view/4003>
- Ullah, S., Akhtar, P. & Zaefarian, G. (2018). Dealing with endogeneity bias: The generalized method of moments (GMM) for panel data. *Industrial Marketing Management*, 71, 69–78. doi: 10.1016/j.indmarman.2017.11.010
- Valdez, M. E. & Richardson, J. (2013). Institutional determinants of macro-level entrepreneurship. *Entrepreneurship Theory and Practice*, 37(5), 1149–1175. doi: 10.1111/2Fetap.12000
- Verheul, I., Wennekers, S., Audretsch, D. & Thurik, R. (2002). An eclectic theory of entrepreneurship: Policies, institutions and culture. In D. Audretsch, R. Thurik, I. Verheul & S. Wennekers (Eds.), *Entrepreneurship: Determinants and policy in a European-US comparison* (pp. 11–81). Springer. doi: 10.1007/0-306-47556-1_2
- Wennekers, A. R. M., Uhlaner, L. M. & Thurik, R. (2002). Entrepreneurship and its conditions: A macro- perspective. *International Journal of Entrepreneurship Education*, 1(1), 25–64. Retrieved from <http://hdl.handle.net/1765/15876>
- Wennekers, S., Thurik, R., van Stel, A. & Noorderhaven, N. (2007). Uncertainty avoidance and the rate of business ownership across 21 OECD countries, 1976–2004. *Journal of Evolutionary Economics*, 17(2), 133–160. doi: 10.1007/s00191-006-0045-1
- Wennekers, S., Van Stel, A., Carree, M. & Thurik, R. (2010). The relationship between entrepreneurship and economic development: Is it U-shaped? *Foundations and Trends® in Entrepreneurship*, 6(3), 167–237. doi: 10.1561/03000000023
- Wennekers, S., Van Wennekers, A., Thurik, R. & Reynolds, P. (2005). Nascent entrepreneurship and the level of economic development. *Small Business Economics*, 24(3), 293–309. doi: 10.1007/s11187-005-1994-8
- Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1), 25–51. doi: 10.1016/j.jeconom.2004.02.005
- Windmeijer, F., Farbmacher, H., Davies, N. & Davey Smith, G. (2019). On the use of the lasso for instrumental variables estimation with some invalid instruments. *Journal of the American Statistical Association*, 114(527), 1339–1350. doi: 10.1080/01621459.2018.1498346
- Wintoki, M. B., Linck, J. S. & Netter, J. M. (2012). Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics*, 105(3), 581–606. doi: 10.1016/j.jfineco.2012.03.005
- Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach*. Nelson Education. Retrieved from <https://www.cengage.uk/shop/isbn/9781337558860>
- Zaefarian, G., Kadile, V., Henneberg, S. C. & Leischnig, A. (2017). Endogeneity bias in marketing research: Problem, causes and remedies. *Industrial Marketing Management*, 65, 39–46. doi: 10.1016/j.indmarman.2017.05.006