PERSISTENCE OF COOPERATION ON INNOVATION: ECONOMETRIC EVIDENCE FROM PANEL MICRO DATA

DOI: 10.18267/j.pep.536

Martin Srholec[*]

Abstract
Arrangements to cooperate on innovation facilitate access to external sources of knowledge. By using panel data derived from the five waves of Community Innovation Survey in the Czech Republic, we examine whether firms engage in these arrangements persistently or rather revert to other behaviour. Econometric estimates of dynamic random effects and multivariate probit models provide strong support to the thesis of persistence, particularly of linkages with the university sector and suppliers. The results are robust to the initial conditions problem and serial correlation in idiosyncratic errors. Government programmes initiating cooperation on innovation therefore have the potential to induce durable changes in the innovative behaviour of firms.

Keywords: innovation, cooperation, persistence, panel data, Community Innovation Survey, Czech Republic
JEL Classification: O31, O32, C23, C25, L20

1. Introduction

Much has been traditionally written about market *versus* hierarchical modes of how firms access external sources of knowledge (Arora *et al*., 2001 and Ahuja and Katila, 2001), for instance, by licensing of patented technology or direct acquisitions of other firms with complementary resources. Arrangements to cooperate on innovation with other enterprises, research institutes or universities represent another one (De Bondt, 1996; Gulati, 1998 and Sachwald, 1998). But the cooperative mode has been often viewed as an uneasy "hybrid", as something suited for extraordinary circumstances, if firms experience temporary shortages of key resources, during episodes of increased risk and when they capture rents from government support. If this point of view is right, cooperation is a transient mode of accessing external knowledge. If in contrast firms tend to persist in cooperation on innovation, there is a need for a different look at this trait of how firms behave.

Availability of micro data from the Community Innovation Survey (CIS) triggered a growing body of empirical research on the cooperative behaviour of firms in the innovation process (Arora and Gambardella, 1994; Colombo, 1995; Veugelers, 1997; Noteboom, 1999; Tether, 2002; Miotti and Sachwald, 2003; Veugelers and Cassiman,

* Martin Srholec, CERGE-EI, Economics Institute, Academy of Sciences of the Czech Republic, Prague, Czech Republic (martin.srholec@cerge-ei.cz).
I am grateful to the Czech Statistical Office for providing access to the micro data. Financial support from the Czech Science Foundation (GAČR) Project P402/10/2310 on "Innovation, Productivity and Policy: What Can We Learn from Micro Data?" and institutional support from the RVO 67985998 from the Academy of Sciences of the Czech Republic is gratefully acknowledged. All usual caveats apply.

2004 and a long list of others later). Until relatively recently, however, this line of research has been hindered by a lack of longitudinal panel data (Mairesse and Mohnen, 2010). So the vast majority of existing studies rely on cross-sectional evidence, which leaves us in the dark with regards to the dynamics, and severely limits the possibilities to tackle endogeneity problems.

Peters (2009) and Raymond *et al*. (2009 and 2010) are rare examples of studies that used panel CIS data to econometrically examine the general question whether firms persist to innovate. Belderbos *et al*. (2004) used lagged predictors to study the cooperative behaviour of firms in innovation on the base of CIS data merged from two periods, but did not explicitly consider the persistence thesis. Only the recent paper by Belderbos *et al.* (2012) examined using this kind of data whether firms cooperate on innovation persistently.

The aim of the paper is to help in filling this gap. Using panel CIS data derived from five waves of CIS in the Czech Republic, we study the cooperative behaviour of firms in a dynamic econometric framework, where the thesis of persistence is represented by a lagged dependent variable. Do firms access external sources of knowledge through cooperative innovation projects on a continuous basis or rather give away to other modes of behaviour as time goes by? Does the propensity to persistence differ by the type of partner for cooperation? And why is the persistence problem important for innovation policy?

The paper proceeds as follows. Section 2 debates reasons why firms do (or not) persist in cooperation on innovation. Section 3 introduces the panel data and explores the descriptive evidence. Section 4 presents results of the dynamic random effects probit model, which considers the general propensity of firms to cooperate. Section 5 further develops the analysis by looking more closely at the different types of partners for cooperation, which requires the multivariate probit model. Section 6 contains concluding remarks and develops policy recommendations.

2. Why Should (or not) Firms Persist in Cooperation on Innovation?

Several interrelated sources of persistence in the cooperative behaviour of firms in the innovation process can be identified. First and perhaps the most obvious reason is that there are sunk costs (Sutton, 1991). If resources required to establish the cooperative link, once spend, cannot be easily used for other purpose, they induce entry and eventually exit barriers. Sometimes cooperation requires a large initial investment in terms of search costs for relevant partners, training costs, organizational changes, trust building between the prospective partners and finding a way to bridge cognitive barriers. Once the relationship is established, however, maintaining the link becomes relatively less demanding.

Furthermore, there are dynamic economies of scale. As the firm accumulates resources, capabilities and experience to innovate in the cooperative manner, the difficulty of using this mode of accessing external knowledge has a tendency to decline over time. Economists often call this process learning-by-doing (Arrow, 1962). But there is also conscious cooperative capability building, because joint innovation projects require specific resources, managerial qualities and “open” sentiment, which do not develop easily (Hildrum, 2010). Accumulating “know-who” about the partners, the element of knowledge in terms of Lundvall and Johnson (1994), is another symptom of this process. Other reasons for continued cooperation that have been cited in the literature include network embeddedness (Granovetter, 1984), firm’s reliance on routines (Nelson and Winter, 1982) and the existence of relational rents (Dyer and Singh, 1998).

From this follows the "persistence hypothesis", according to which cooperation is a durable trait in the innovative behaviour of firms. Firms are expected to commit resources to cooperation on innovation on a continuous basis, and therefore this mode of accessing external knowledge tends to persevere over time. Arguments in favour of the persistence thesis could be also labelled under the heading of "accumulation" premise, because they share the emphasis on continuity, feedbacks, inertia and hence path dependence of how firms behave.

According to the transaction cost theory, however, firms should be extremely reluctant to cooperate on innovation. Since innovation is a highly uncertain venture (Dosi, 1988), budgets, running and outcomes of joint innovation projects are very hard to foresee, and therefore property rights cannot be clearly defined *ex-ante*. If complete contracts cannot be written, which is exactly the case here, opportunistic behaviour is likely to occur, because the partners cannot prevent the use of knowledge pooled, utilized and generated in the joint project outside of the contract (De Bondt, 1996). Cooperation on innovation should be the option of a last resort.

From this follows the rival "transience hypothesis", which postulates that cooperation is a temporary solution. Firms are expected to cooperate on innovation in extraordinary circumstances, for instance if they experience temporary shortages of key resources that cannot be obtained otherwise, during episodes of increased risk or in order to capture rents from public support to cooperative projects. But after these temporary incentives expire they are supposed to give away to the traditional modes of accessing external sources of knowledge. According to this view, cooperation on innovation is a nuisance, an uneasy solution, a curiosity in how firms behave.

Most of the existing papers using panel methods on data from CIS (or very similar) surveys focus on the general questions whether firms persist on innovation, not specifically whether they cooperate. Peters (2009) and Raymond *et al*. (2009) found support for persistence of both innovation inputs and outputs. Raymond *et al.* (2010) using a novel method to treat the initial conditions problem only confirmed true persistence in the propensity of firms to innovate in the high-tech sector, while other results came out inconclusive. Earlier studies by Van Leeuwen (2002), Duguet and Monjon (2004) and Rogers (2004) found support for the persistence thesis but because of data constraints could not control for unobserved individual heterogeneity.

Belderbos *et al*. (2012) is the only attempt to examine the persistence of cooperation on innovation in a longitudinal panel data design so far. Since the panel-level variance component did not come out significant, they gave more weight to presenting estimates of a pooled multivariate model, but irrespectively of the method, their results are overwhelmingly in favour of the persistence thesis. Overall, therefore, the existing evidence is largely in favour of persistence in cooperation on innovation and in the innovative behaviour of firms at large. But the literature on this topic remains thin and limited to evidence from a small number of countries.

3. Data

The empirical analysis is based on micro data from CIS conducted by the Czech Statistical Office, which are fully harmonized with methodology of the Oslo Manual (OECD, 1997 and 2005). CIS targets enterprises with at least 10 employees; the data has been collected by a census of large firms with at least 250 employees and a sample survey of smaller firms stratified by industry, size categories and in the last three surveys by NUTS3 regions. Since answering compulsory by law, response rates ranged between 63 to 83%, which is quite high for this kind of data.

To create a panel dataset we merged five consecutive waves of CIS with reference periods 1999–2001, 2002–2003, 2004–2006, 2006–2008 and 2008–2010. From this follows, that the periods were three-year, only except of the second survey, and there is one-year overlap between the last three surveys, which one needs to keep in mind, as this is a potential source of bias. CIS collects data for firms with principal activity classified in industry and market services (10–74), but because some sectors have been covered irregularly, we exclude from the dataset the sectors of construction (45), repair, wholesale and retail trade (50–52), hotels and restaurants (55), real estate (70) and renting services (71); NACE, rev. 1.1 codes are denoted in the brackets.

As a consequence of the random stratified sampling, which draws a somewhat different pool of respondents from the targeted population for each wave of the survey, many firms appear in the data only once. But to study the dynamics of their behaviour we need to observe them repeatedly. Hence, we rely on an unbalanced panel of about 4,000 firms, which are present in at least two consecutive surveys, and which covers 31% of the total dataset. It is well acknowledged that this is a noticeable reduction, but for similar datasets used in most other countries the loss of observations is in fact much worse.[1]

Table 1 provides definitions of the variables. At the centre of our interest there are the variables for cooperation, which are derived from the set of questions on whether the firm cooperated on any of its innovation activities with other organizations; namely with i) suppliers; ii) customers; iii) competitors; iv) consultants, commercial labs, or private research institutes; v) government and public research institutes; and vi) universities or other higher education institutions.[2] From this follows a set of six CO type dummy variables with value 1 for firms that engaged in cooperation with the respective type of partner and an overall CO dummy with value 1 for firms that cooperated with at least one of them.

Furthermore, there is a battery of variables accounting for other characteristics of the firms. R&D that captures the internal technological resources is as a dummy variable with value 1 if the firm continuously engaged in research and experimental development activity. GP variable on whether the firm is part of a (domestic or foreign) group, *AGE* given by the (log of) years since registration and *SIZE* measured by the (log of) initial employment account for structural features. Sectoral differences are controlled for by a set of industry dummies denoted by *INDUSTRY* broadly following 2-digit level of *NACE*, rev. 1.1.[3] Finally, *PERIOD* dummies are used to account for the presence of common shocks.

1 Not surprisingly, given the sampling methodology, there is a bias of the sample in favour of large firms, because collecting data for them by a census naturally boosts their chances to appear in the surveys repeatedly. Nevertheless, the sectoral composition appears very similar to the overall dataset, not raising concerns of selectivity along these lines.

2 Another type of partner are other firms within the respondent's group. But because only firms affiliated to a group can answer affirmatively, and because internal cooperation within the group is likely to follow different rules, we only consider cooperation with the external partners.

3 Because of a small number of observations the following NACE, rev. 1.1 categories must have been combined together: i) 10, 11, 12; ii) 15, 16; iii) 18, 19; iv) 23, 24; v) 30, 31; vi) 36, 37; vii) 60, 61, 62; and 65, 66, 67.

Table 1 | Definition of the Variables

CO	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with an external, *i.e*. non-affiliated, partner
COsupp	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with suppliers of equipment, materials, components, or software
COcust	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with clients or customers
COcomp	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with competitors or other enterprises in the same sector
COprivlab	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with consultants, commercial labs, or private research institutes
COgovlab	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with government or public research institutes
COuni	Dummy variable with value 1 if the firm had a cooperation arrangement on innovation with universities or other higher education institutions
R&D	Dummy variable with value 1 if the firm reported to continuously engage in intramural research and experimental development activity
GP	Dummy variable with value 1 if the firm was affiliated to a group
AGE	Log of the number of years since the firms has been recorded in the business register
SIZE	Log of the number of employees in the initial year of the reference period
INDUSTRY	Dummy variable with value 1 if the principal activity of the firm was classified in the respective 2-digit NACE, rev. 1.1 industry
PERIOD	Dummy variable with value 1 if the observation was recorded in the respective wave of the survey

Source: Own computation based on CIS micro data from the Czech Statistical Office

Table 2 provides the descriptive overview. After omitting observations with incomplete records, we arrive to a panel of 3,942 firms with 11,142 observations, for about a half of which we have data in two consecutive periods only, hence this is a highly unbalanced panel. About one fourth of the firms reported to cooperate on innovation. Other enterprises, namely suppliers, are the most popular, while public research institutes are the least frequent partners. But this should not be necessarily interpreted as a sign of weak national research infrastructure, because similar proportions are common in advanced countries too.

Nevertheless, for the purpose of this study the most important information is the extent of variation of the variables "between" firms in a given period as compared to the variation "within" the same firm over time. All of the variables show less variation within than between, thus signalling persistence, although for the cooperation variables the difference is relatively small, so the descriptive evidence is inconclusive. To find out whether there is true support for the persistence thesis or whether the cooperative behaviour of firms is actually driven by other factors is the main purpose of the econometric analysis.

Table 2 | Descriptive Statistics

	Mean	Overall st. dev.	Between st. dev.	Within st. dev.	Number of firms	Number of observations
CO	0.25	0.43	0.33	0.27	3,942	11,142
COsupp	0.19	0.39	0.28	0.26	3,942	11,142
COcust	0.15	0.36	0.26	0.24	3,942	11,142
COcomp	0.08	0.28	0.19	0.20	3,942	11,142
COprivlab	0.13	0.33	0.24	0.22	3,942	11,142
COgovlab	0.06	0.24	0.17	0.16	3,942	11,142
COuni	0.12	0.33	0.25	0.20	3,942	11,142
R&D	0.24	0.43	0.35	0.23	3,942	11,142
GP	0.46	0.50	0.45	0.22	3,942	11,142
AGE	2.41	0.52	0.49	0.23	3,942	11,142
SIZE	4.93	1.45	1.38	0.36	3,942	11,142

4. Random Effects Probit Model

The aim of the model is to investigate the persistence of cooperative behaviour of firms in the innovation process, whether they cooperate repeatedly or rather tend to switch between the cooperative and other modes of behaviour. Since the dependent variable is binary, we estimate random effects probit model, of which the familiar variance components notation is as follows:

$$CO_{it} = 1 \text{ if } \alpha CO_{it-\tau} + \beta x_{it} + \delta_i + \varepsilon_{it} > 0 \text{ and } 0 \text{ otherwise} \qquad (1)$$

where i denotes the firm ($i = 1, ..., N$), t is time ($t = 1, ..., T\text{-}\tau$), τ is the time lag; the current decision to cooperate CO_{it} is the function of past cooperation $CO_{it-\tau}$, other observable characteristics x_{it}, unobserved time-invariant individual effects δ_i that are conventionally assumed to be independent and identically sampled from a normal distribution with expected zero mean and variance σ_δ^2, and other unobserved time-variant effects ε_{it} that are assumed to be independent and identically Gaussian distributed with mean zero and variance σ_ε^2 ,= 1.

At the centre of our interest is the estimate of state dependence given by α, as a significantly positive coefficient signifies the persistence of cooperation, while a significantly negative result lends support to the transience thesis. Other observable characteristics that are likely to affect the outcome are represented by the vector of $x_{it} \in (R\&D_{it}, GP_{it}, AGE_{it}, SIZE_{it}, INDUSTRY_{it}, PERIOD_{it})$. But other potentially relevant characteristics are not observed. If these were not properly accounted for, the lagged dependent variable could seem to determine the current outcome because of picking up their effects. Heckman (1981a) calls this "spurious state dependence", which should not be confused with the "true state dependence". For this purpose, the latent individual effect δ_i is included to control for the time-invariant unobserved attributes, such as latent capabilities, perceptions and risk profiles of firms.

Table 3 gives the results. The estimated models include the battery of industry dummies to account for sectoral differences and the set of time dummies to control for cross-sectional dependence. But for the sake of brevity the respective coefficients are not reported, because these variables are primarily used to control for the context-specific effects and do not warrant closer attention in this study by themselves. Following the method of Naylor and Smith (1982), adaptive Gauss–Hermite integration with 12 quadrature points is used in the estimates, however, the results are highly robust to the choice of integration method and points. [4]

Table 3 | Results of Random Effects Probit Model for CO_{it} as the Dependent Variable

	(1)		(2)		(3)	
	$\tau = 1$		$\tau = 1, 2$		$\tau = 1, 2, 3$	
CO_{it-1}	0.89	(0.05)***	0.85	(0.07)***	1.05	(0.09)***
CO_{it-2}	..		0.37	(0.07)***	0.41	(0.11)***
CO_{it-3}	..		..		0.32	(0.12)***
$R\&D_{it}$	1.15	(0.05)***	1.17	(0.08)***	1.13	(0.12)***
GP_{it}	0.20	(0.04)***	0.16	(0.06)**	0.29	(0.10)***
AGE_{it}	–0.02	(0.05)	–0.04	(0.09)	–0.24	(0.18)
$SIZE_{it}$	0.15	(0.02)***	0.14	(0.03)***	0.07	(0.05)
Constant	–2.36	(0.18)***	–2.20	(0.32)***	–1.56	(0.60)**
Industry dummies	Yes		Yes		Yes	
Period dummies	Yes		Yes		Yes	
Number of observations	7,147		3,152		1,363	
Number of firms	3,942		1,789		923	
ρ	0.03	(0.04)	0.05	(0.06)	0.04	(0.11)
Wald χ^2	1,735.33***		786.88***		271.41***	
Log–likelihood	–2,832.39		–1,347.55		–555.22	

Note: Standard errors in brackets; ***, **, and * indicate significance at the 1, 5, and 10 per cent level.

In the first column, we present a benchmark model with $\tau = 1$, thus the dynamics is first order only. CO_{it-1} comes out with a positive and highly statistically significant coefficient, which indicates the persistence of cooperation. After controlling for the other relevant effects, firms that cooperated on innovation in the past are confirmed to be considerably more prone to cooperate in the present. In terms of marginal effects, derived at the mean of the other covariates, the past cooperation increases the likelihood of present cooperation by

4 For details on the underlying maximum likelihood procedures, methods and formulas see Stata (2009, pp. 408–427).

25.9 percentage points.[5] Arguably, this result signifies a sizeable degree of persistence in the cooperative behaviour of firms.

Next, in the second column, there are results of a model with $\tau = 1$ and 2, thus the dependent variable is lagged by one and two reference periods of the survey. As a result, the sample consists only of firms observed in at least three consecutive periods, and hence the number of observations drops down noticeably. Again, the tendency to persistence is strongly confirmed. In terms of the marginal effects, holding all other covariates at the mean, the probability of cooperation in the current period is estimated to increase by 29.5 and 12.9 percentage points with the past cooperation lagged by one and two periods, respectively. And both of the lagged coefficients remain highly statistically significant.

Finally, we add a third lag of the dependent variable, thus $\tau = 1$, 2 and 3, implying a model with third order dynamics in the last column. Estimating this specification requires the firms to be observed in at least four consecutive periods, which is quite stringent for data based on stratified sampling, so the sample shrinks to less than a quarter of the original. Not surprisingly, the resulting sample has biased composition; larger, grouped and R&D performing firms are overrepresented, hence the results are strictly speaking not comparable to the benchmark. Anyhow, the persistence thesis holds firmly. All three of the lagged coefficients come out highly significant and the corresponding marginal effects are 38.0, 14.7 and 11.5 percentage points, respectively. Hence, if combined together, this result implies that a firm cooperating three times in a row is estimated to have by 64.2 percentage points higher propensity to cooperate once again.[6]

As far as the other observable characteristics are concerned, the results are broadly in line with the previous empirical research on this topic (Miotti and Sachwald, 2003; Veugelers and Cassiman, 2004; Srholec, 2009). Most of them come out with significantly positive coefficients, except only for the AGE_t the firm, which does not seem to make a difference. $R\&D_{it}$ represents not only the ability of firms to generate new knowledge, but as pointed out by Cohen and Levinthal (1990), also their capacity to absorb knowledge from outside, so this is not surprisingly a highly relevant covariate. GP_{it} boosts the odds of cooperation, because affiliated firms benefit from linkages to the potential partners developed by other members of the group. $SIZE_{it}$ representing scale advantages is essential to control for, which is confirmed by the results, except for the final estimate, possibly because of the bias of this sample in favour of large firms.

So far we have not discussed the unobserved effects. At the bottom of the table there is reported the estimated parameter $\rho = \dfrac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\varepsilon^2}$, which gives the proportion of the total variance contributed by the individual variance component. If ρ is close to zero, in other words, the latent individual effect δ_i is unimportant. Since a likelihood ratio test whether

5 A marginal effect generally refers to the percentage change in the probability of a success in response to one percentage change in the explanatory variable, holding all other variables at some fixed values. Specifically for binary explanatory variables the marginal effect is computed for changing their value from zero to one. To derive the marginal effect in this dynamic random effects probit model, we must assume that the latent individual effect is zero.

6 Unfortunately, the data is insufficient to compute a model with $\tau = 1$, 2, 3 and 4, thus allowing for dynamics of fourth order, because this requires a balanced panel, for which there is a sample of 412 firms only, the composition of which is heavily biased and which is insufficient for deriving a reliable estimate.

ρ is different from zero comes out statistically insignificant at the conventional levels and the proportion accounted by the unobserved individual variance is small, at least in this sample specification this effect does not seem to be a serious matter for concern. [7]

As already mentioned above, there can be a bias in favour of detecting persistence, because of the one-year overlap between the reference periods of the last three surveys, *i.e.* 2004–2006, 2006–2008 and 2008–2010. Naturally, if a firm cooperated in the common year, there is an affirmative answer in both periods. Thus, we repeated the estimates on data without the last but one survey, *i.e.* 2006–2008. The results are fairly robust in this respect. In terms of the marginal effects, all else constant at the mean, the re-estimated degree of persistence is 21.9 percentage points in the benchmark model and 22.6 and 7.7 percentage points for the two consecutive lags in the second model; all highly statistically significant. If the third model is run on the reduced dataset, which implies using a balanced panel, the marginal effects are estimated to be 29.4, 7.3 and 4.9 percentage points for the three respective lags; however, the data is insufficient to compute reliable estimates of standard errors. Overall, hence, the persistence thesis holds.

Endogeneity problems are pandemic in the literature on cooperation on innovation, because the vast majority of existing papers is based on cross-sectional evidence. A major caveat of these studies is that the potential simultaneity bias can be treated only by using instrumental variables, which are extremely hard to find. Even though the dynamic panel data model estimated above requires that the effects of x_{it} are taken to be strictly exogenous, because this is necessary to derive the likelihood of observing a given series of outcomes as the product of individual likelihoods, the lagged dependent variable $CO_{it\text{-}\tau}$, which is the focal point, is by construction allowed to be correlated with δ_i and with lagged ε_{it}. Hence, in this respect we operate on noticeably safer ground than the cross-sectional papers.

Yet as often in econometrics there is a catch. If ε_{it} is serially correlated, $CO_{it\text{-}\tau}$ may also be correlated with the current ε_{it}, reincarnating the endogeneity concerns, so for the presented results to be consistent, we need to assume that this is not the case. Moreover, we have ignored the potential problem of initial conditions described by Heckman (1981b) that besets estimating dynamic probit models. If the serial correlation of ε_{it} is positive, there is an upward bias of α, hence in favour of detecting persistence, and *vice versa*. Likewise, if the initial conditions are positively correlated with the δ_i, the method of estimation assuming their independence tends to overstate the degree of persistence α. Hence, in the next step we test robustness of the results to relaxing these restrictions.

Heckman (1981b) proposed a solution of the initial conditions problem which involves approximating the conditional distribution of the initial value of the dependent variable:

$$CO_{i1} = 1 \text{ } if \text{ } \pi z_{i1} + \eta\delta_{\mathrm{i}} + \varepsilon_{i1} > 0 \text{ and } 0 \text{ otherwise} \tag{2}$$

where z_{i1} is a vector of exogenous covariates which includes x_{i1} and additional (preferably presample) instrumental variables in the initial period, δ_i and ε_{i1} are independent of each other and ε_{i1} meets the same distributional assumptions as ε_{it}. Equation 1 remains the

7 After all, if the model is estimated by a simple pooled probit model, the results are nearly the same. Note that the insignificance of the individual variance component can be more than anything else the consequence of having a relatively short panel. It will be interesting to see whether this effect becomes more significant in future research based on longer time series.

same, except that now $t = 2, ..., T$, thus there is one less period available to estimate α, as $t = 1$ represents the initial period.

Wooldridge (2005) put forward what has been dubbed the "simple solution" of the initial conditions problem, which is based on a different approximation. His estimator considers the distribution of the conditional individual latent effect δ_i:

$$\delta_i = \mu + \gamma CO_{i1} + \psi \overline{x}_i + \xi_i \tag{3}$$

where $\overline{x}_i = \frac{\sum_{t=2}^{T} x_{it}}{(T-1)}$, *i.e.* the time-means of x_{it} excluding the initial period, and ξ_i are assumed to be independent and identically sampled from a normal distribution with expected zero mean and variance σ_ξ^2.[8] After integrating out ξ_i the likelihood has exactly the same structure as in the random effects probit model in Equation 1; except only that the additional covariates CO_{i1} and $\overline{x}_i$ are included, hence we substitute δ_i from Equation 3 in Equation 1 and use the same estimation procedure. No instruments are required, however, the model is estimated on data from one less period than the Heckman (1981b) solution, which is unattractive for data with a short panel.

Stewart (2007) extended the estimator by Heckman (1981b) to the case of serially correlated errors in which ε_{it} follows a first-order autoregressive process, which can be for the sake of brevity delineated as follows:

$$\varepsilon_{it} = \lambda \varepsilon_{it-1} + \varphi_{it} \text{ and } -1 < \lambda < 1 \tag{4}$$

where λ represents the serial correlation of ε_{it}. Note that the computation involves T-dimensional integrals which complicates the estimation immensely and proves to be highly data demanding.

A major drawback is that these estimators have been developed for balanced panels only. Nevertheless, it is instructive to derive consistent estimates on the sub-sample of balanced data in order to obtain at least a rough idea about the magnitude (and direction) of the bias and hence to get one step closer to determining whether the true state dependence exists. Akay (2011) shows that the Wooldridge's method works well for panels with moderate durations, while the Heckman's approximation is suggested for short panels, but the latter hinges on the instruments. Hence, for comparative purpose, we present results based on both estimators noting that the truth is likely to lie somewhere in between.

Because of data limitations we have to simplify the specification. First, the balanced panel is generated excluding the first period, which almost doubles the sample to 768 firms, as compared to only 412 firms in the full five-period balanced panel. Second, due to the sample size reduction the industry dummies need to be aggregated to 19 categories broadly following two-letter alphabetical classification of *NACE*, rev. 1.1. Finally, the model is restricted to $\tau = 1$, thus the dynamics is first order.

In addition, the Heckman's method and the follow up Stewart's estimator require instruments of the initial value of the dependent variable. Besides x_{ij} the vector z_{ij} therefore contains additional variables recorded in the first period, namely a dummy for the legal form with value 1 if the firm was a joint-stock company and a dummy for state ownership

8 Wooldridge (2005) used the full set of $x_i = (x_{i2}, ..., x_{iT})$ instead of $\overline{x}_i$, but using the time-means gives essentially the same results, while substantially reducing the number of extra covariates to be estimated, hence recent implementations prefer the latter solution.

with value 1 if the firm was state-owned, both of which turned out to be relevant proxies of the pre-sample conditions.[9]

Table 4 | Results of Alternative Estimators for CO_{it} as the Dependent Variable

	(1)		(2)		(3)	
	Wooldridge (2005)		Heckman (1981b)		Stewart (2007)	
CO_{it-1}	0.88	(0.11)***	0.71	(0.11)***	1.19	(0.11)***
$R\&D_{it}$	0.84	(0.15)***	1.24	(0.09)***	1.11	(0.09)***
GP_{it}	0.50	(0.23)**	0.23	(0.08)***	0.20	(0.07)***
AGE_{it}	0.71	(2.14)	−0.10	(0.13)	−0.09	(0.11)
$SIZE_{it}$	−0.41	(030)	0.22	(0.04)***	0.16	(0.04)***
Mean($R\&D_i$)	0.73	(0.23)***	..		..	
Mean(GP_i)	−0.28	(0.26)	..		..	
Mean(AGE_i)	−0.80	(1.98)	..		..	
Mean($SIZE_i$)	0.56	(0.30)*	..		..	
Constant	−2.43	(0.73)***	−2.03	(0.43)***	−1.76	(0.37)***
Industry dummies	Yes		Yes		Yes	
Period dummies	Yes		Yes		Yes	
Number of observations	2,304		3,072		3,072	
Number of firms	768		768		768	
ρ	0.16	(0.11)*	0.18	(0.07)**	0.11	(0.05)**
CO_1	0.39	(0.13)***	..		..	
η	..		0.35	(0.29)	0.22	(0.37)
λ	..		..		−0.30	(0.05)***
Wald χ^2	335.04***		549.40***		704.61***	
Log-likelihood	−644.29		−1,339.77		−1,331.00	

Note: Standard errors in brackets; ***, **, and * indicate significance at the 1, 5, and 10 per cent level.

Table 4 provides the alternative estimates. Results of the methods by Wooldridge, Heckman and Stewart are presented in the first, second and thirds columns, respectively. Exogeneity of the initial conditions in the standard random effects model can be seen as imposing that the coefficients of CO_{i1} and η equal to zero. The hypothesis that CO_{i1} is zero is strongly rejected which implies that indeed there was a bias. The estimate of η is not statistically significant at the conventional levels, thus less precise, perhaps because the instruments are not strong enough. Nevertheless, the estimated magnitude of the bias is relatively similar, *i.e.* 0.39, 0.35 and 0.22, even though the first is based on a sample reduced by one more period than the latter two.

9 Both have a highly significant positive impact (individually as well as jointly) on the probability of cooperation in the initial period probit equation. Other variables, for instance, the mode of entry, were also tested as potential instruments but did not come out empirically relevant.

In the third column, the Stewart's estimator further allows for the serial correlation in ε_{it}. The estimate of λ is highly statistically significant, thus the hypothesis that ε_{it} are serially independent is rejected. Clearly, the negative sign, which implies that successive ε_{it} are negatively correlated, strikes as a peculiar result. One possibility is that the transitory disturbances reflect the intertemporal characteristics of research process, on which hinges the cooperative behaviour, in the sense that the occurrence of great ideas is unstable over time, there are outbursts of activity followed by periods of tranquillity. Another possibility is that due to the three-year length of reference periods, the data tend to capture only peaks and bottoms in innovation (as wells business and political at large) life cycles, not the trajectory in between, as the results of which there seem to be positive shocks followed by the negative ones and *vice versa*.

Most importantly, the estimated coefficient of the lagged dependent variable CO_{it-1} confirms large and highly statistically significant first-order state dependence; hence the main conclusions turn out to be remarkably robust. The magnitude is 0.88 using the Wooldridge's estimator, which is as expected a bit smaller than 1.06, if the standard random effects probit is estimated on the same sample. The magnitude is 0.71 using the Heckman's method, but 1.19 using the Stewart's estimator, which compares to 0.77 using the standard random effects estimator. From this follows that accounting for the initial conditions in one way of the other slightly reduces the estimate of α, while allowing for the serial correlation in e_{it} quite dramatically boosts the estimated degree of persistence, hence these assumptions work in opposite directions, but the dominant bias is toward zero. Judging from this limited evidence, therefore, the true degree of persistence seems to be understated by the standard random effects probit estimator.

5. Multivariate Probit Model

Following Belderbos *et al*. (2012) we examine the heterogeneity of partners, namely whether firms tend to stick to cooperation with the same type of partner or rather switch between the different types. Unlike in the previous section, we examine a situation with more than one possible outcome. If we estimated separate univariate models, we would ignore the possibility that the decisions of firms to cooperate with the different types of partners might be correlated with each other. To allow for their interdependence, we delineate a system of multiple equations, which considers probability of the different outcomes jointly in the following variance components model:

$$CO_{it(j)} = 1 \text{ if } \alpha_{(j)} CO_{it-\tau(j)} + \beta_{(j)} x_{it(j)} + \varepsilon_{it(j)} > 0 \text{ and 0 otherwise} \tag{5}$$

where j refers to the six types of partners, $j = 1 \ldots 6 = COsupp_{it}$, $COcust_{it}$, $COcomp_{it}$, $COprivlab_{it}$, $COgovlab_{it}$ and $COuni_{it}$; $\varepsilon_{it(j)}$ indicates six residuals sampled from the multivariate normal distribution that are allowed to be correlated with each other, so there is variance-covariance matrix V with values of 1 on the leading diagonal and these correlations $\omega_{kj} = \omega_{jk}$ as off-diagonal elements. Hence, a statistical test for $\omega = 0$ can be estimated that examines the hypothesis of interdependence between the different types. If the correlation is positive, the propensity of cooperation with the respective type of partner increases simultaneously with the other one, perhaps because they are complementary. But the correlation turns out negative if there are trade-offs between them.

Besides including the lagged dependent variables by the type of partner on the right hand side, the other covariates x_{it} remain the same. The dynamics is second order, thus $\tau = 1$ and 2, as the data is insufficient to estimate a three-lag model. The latent individual effect δ_i is not estimated, as including this component proved to be unnecessary according to the random effects probit results, and because omitting this parameter greatly simplifies the computation. Hence, more specifically, we use pooled multivariate probit. But the estimated standard errors are clustered at the individual level to account for the fact that the observations are not independent within. [10]

Table 5 shows the results. The most lasting partnerships are being developed with universities followed by linkages along the value chain with suppliers and customers, as the persistence thesis holds in both lags. Somewhat less durable appear the cooperative deals with private research sector, public research institutes and competitors, for which the first lag comes out statistically significant, but the second lag does not. For the competitors this is well in line with expectations, because of the risk of opportunistic behaviour, but a suitable interpretation for the lower persistence of cooperation with the research organizations regardless of ownership is not obvious, especially if contrasted with the university sector. If the joint significance of the first and second lags is tested, *i.e*. the null hypotheis $CO_{it\text{-}1(j)} = CO_{it\text{-}2(j)} = 0$ in the respective j-equation, the test statistics rejects the null at 1% level for all types of partners, so there is a great deal of persistence across the broad, but noticeably far more for the universities and suppliers.

Overall, the results are in line with the conclusions of Belderbos *et al.* (2012) based on the Dutch CIS data that are overwhelmingly in favour of the persistence thesis with the clarification that they found stronger persistence in linkages with customers than suppliers. Arguably, this can be attributed to structural differences between the Dutch and Czech economy, namely that most firms in the latter are well behind the global market leaders, hence their innovation strategies tend to be relatively more driven by suppliers.

Another key result of the multivariate model can be found in the lower part of the table, where correlations between the residuals are reported. All of them are significantly positive, which indicates that the firms combine the different types of partners. Strictly speaking, this may not necessarily be due to their complementary nature, but due to omitted factors affecting them jointly. But we side with the former interpretation, because the relationships come out very strong. In this respect the results are in accord with Belderbos *et al*. (2012), even though they only considered links between different types of business partners, and support the burgeoning literature on systemic interactions in the innovation process reviewed, for example, by Soete *et al*. (2010).

As far as the lagged coefficients of other lagged cooperation types are concerned, which account for the possibility that past cooperation with the respective type affects the odds of engaging the current type, the results are less conclusive. All of at least weakly statistically significant cross-type coefficients are positive, indicating chronological compatibility of the different partners, thus confirming the findings of Belderbos *et al*. (2012), but there is not an easily recognizable pattern, except perhaps that there is a weak evidence that the firms tend to switch to private research organization from suppliers and to public research institutes from cooperating with other partners.

10 For details on the simulated maximum likelihood procedure used to estimate this model see Cappellari and Jenkins (2003).

Table 5 | Results of Multivariate Probit Model for CO_t by the Type of Partner as the Dependent Variable

	(1)		(2)		(3)		(4)		(5)		(6)	
	COsupp$_{it}$		***COcust$_{it}$***		***COcomp$_{it}$***		***COprivlab$_{it}$***		***COgovlab$_{it}$***		***COuni$_{it}$***	
COsupp$_{it-1}$	0.77	(0.09)***	0.15	(0.09)	0.13	(0.11)	0.21	(0.09)**	0.11	(0.10)	0.06	(0.10)
COsupp$_{it-2}$	0.22	(0.10)**	0.14	(0.10)	0.13	(0.11)	0.27	(0.11)**	0.22	(0.12)*	0.05	(0.11)
COcust$_{it-2}$	0.01	(0.09)	0.54	(0.10)***	−0.05	(0.12)	−0.14	(0.10)	−0.08	(0.12)	−0.07	(0.11)
COcust$_{it-2}$	−0.07	(0.10)	0.23	(0.11)**	0.01	(0.12)	−0.08	(0.12)	−0.07	(0.13)	0.12	(0.12)
COcomp$_{it-1}$	−0.01	(0.10)	0.03	(0.11)	0.75	(0.12)***	−0.08	(0.11)	0.23	(0.12)*	0.13	(0.11)
COcomp$_{it-2}$	−0.05	(0.11)	−0.11	(0.11)	−0.04	(0.12)	−0.10	(0.13)	−0.03	(0.13)	−0.19	(0.12)
COprivlab$_{it-1}$	0.14	(0.09)	0.01	(0.10)	0.21	(0.10)**	0.76	(0.09)***	0.18	(0.11)*	0.15	(0.11)
COprivlab$_{it-2}$	0.04	(0.11)	0.03	(0.10)	0.03	(0.11)	0.11	(0.12)	−0.01	(0.12)	0.04	(0.12)
COgovlab$_{it-1}$	−0.03	(0.11)	−0.05	(0.11)	−0.04	(0.12)	0.13	(0.11)	0.85	(0.11)***	0.12	(0.11)
COgovlab$_{it-2}$	0.06	(0.11)	−0.10	(0.12)	−0.11	(0.13)	0.03	(0.13)	0.08	(0.13)	−0.04	(0.13)
COuni$_{it-1}$	0.12	(0.10)	0.21	(0.10)**	−0.01	(0.11)	0.07	(0.10)	0.22	(0.11)**	0.88	(0.10)***
COuni$_{it-2}$	0.08	(0.11)	−0.09	(0.11)	0.07	(0.12)	0.04	(0.11)	0.06	(0.12)	0.41	(0.11)***
R&D$_{it}$	0.84	(0.07)***	0.90	(0.07)***	0.84	(0.08)***	0.96	(0.07)***	0.82	(0.09)***	0.96	(0.07)***
GROUP$_{it}$	0.26	(0.06)***	0.16	(0.07)**	0.08	(0.08)	0.21	(0.07)***	−0.12	(0.08)	0.01	(0.07)
AGE$_{it}$	0.05	(0.09)	0.03	(0.10)	0.01	(0.11)	−0.06	(0.10)	0.01	(0.11)	−0.05	(0.11)
SIZE$_{it}$	0.15	(0.03)***	0.12	(0.03)***	0.11	(0.03)***	0.09	(0.03)***	0.10	(0.03)***	0.17	(0.03)***
Constant	−2.12	(0.31)***	−2.06	(0.35)***	−2.16	(0.40)***	−2.06	(0.34)***	−2.56	(0.40)***	−2.80	(0.39)***
Industry dummies	Yes		Yes		Yes		Yes		Yes		Yes	
Period dummies	Yes		Yes		Yes		Yes		Yes		Yes	
ω_{2j}	0.83	(0.02)***										
ω_{3j}	0.70	(0.03)***	0.76	(0.02)***								
ω_{4j}	0.74	(0.02)***	0.70	(0.02)***	0.70	(0.03)***						
ω_{5j}	0.63	(0.03)***	0.64	(0.03)***	0.68	(0.03)***	0.67	(0.03)***				
ω_{6j}	0.66	(0.03)***	0.68	(0.02)***	0.67	(0.03)***	0.72	(0.03)***	0.75	(0.03)***		
Number of observations	3,152											
Number of firms	1,789											
Wald χ^2	2,752.63***											
Log-likelihood	−4,635.34											

Note: Clustered standard errors in brackets; ***, **, and * indicate significance at the 1, 5, and 10 per cent level.

Finally, the other covariates included in the model convey qualitatively similar picture across the types with the exception that being part of a group significantly boosts the odds of cooperation with partners in the private sector only, which is feasible given the fact that this implies a business connection.

6. Conclusions

Using longitudinal panel data derived from five waves of CIS in the Czech Republic, we found a credible support for the thesis that cooperation is a persistent, not only a transient, trait in the innovative behaviour of firms, and that this holds particularly for their linkages with the university sector and suppliers. Cooperation on innovation turns out to be an established organizational form, to which firms commit resources repeatedly, not an inconvenient "hybrid", from which firms revert to the traditionally researched market or hierarchical modes of accessing external sources of knowledge.

So what does this mean for innovation policy? First and foremost, this is good news for policy makers, because if the persistence thesis holds, government programmes stimulating cooperation on innovation have a potential to induce lasting changes in the behaviour of firms. Generally speaking, the results provide support to the line of policy interventions directly aimed at facilitating learning between organizations advocated by Lundvall and Borrás (2005), mitigating "network failures" in terms of Woolthuis, *et al.* (2005), and correcting the "systemic failures" highlighted by Chaminade and Edquist (2006), which jointly point to the problem of deficient linkages, complementarities and cross-fertilization of ideas between parts of the system.

At the same time, however, the tendency to persistence represents a challenge for policies of this kind, because it is by principle not efficient to support continuation of behaviour, which tends to persist. From this follows that the prime purpose of the government support programmes should be to kick-start cooperation, to provide the initial "big push" incentive for the partners to get together, but then support the particular linkage just long enough to take root. Needless to say, if the subsidized collaborative project is successful, this provides the best positive feedback for the firm to keep on doing so. Hence, the key practical questions that administrators of these programs should scratch their heads about are exactly how long is just long enough and how to avoid repeatedly handing out subsidies to bundles of the same parties.

Admittedly, this is a topical problem, as governments increasingly devote resources to policy schemes stimulating cooperation on innovation. For instance, the Technology Agency of the Czech Republic (TACR) has launched a large scale "Competence Centres" programme specifically aimed at promoting longstanding collaboration between the private and public sectors, which has a combined budget of CZK 6 billion over 2012–2019, and which is expected to support about 35 projects each combining at least three firms and one public research organization over the period of six years or longer. It remains to be seen whether this programme is going to be efficient in the sense mentioned above and whether the results will be evaluated with these issues in mind.

Besides the econometric issues addressed above, a major limitation that needs to be mentioned is that ideally we should trace the individual cooperation projects (we should use project level data). Admittedly, the data in hand may refer to different projects conducted by a firm over time. Hence, whatever relationship we find therefore refers to the behavioural trait of the firm as a whole, not to persistence of the specific project or whether there persists a link with the same individual partner, and the results should be interpreted accordingly. Understanding of particular relationships, which goes beyond the scope of this paper, needs to be deepened by case-study analysis of the individual cases.

References

Ahuja, G., Katila, R. (2001). Technological Acquisitions and the Innovation Performance of Acquiring Firms: A Longitudinal Study. *Strategic Management Journal*, 22(3), 197–220. DOI: 10.1002/smj.157.

Arora, A., Gambardella, A. (1994). Evaluating Technological Information and Utilizing It. *Journal of Economic Behaviour and Organization*, 24(1), 91–114, DOI: 10.1016/0167-2681(94)90055-8.

Arora, A., Fosfuri, A., Gambardella, A. (2001). *Markets for Technology: The Economics of Innovation and Corporate Strategy*. Cambridge (MA): MIT Press.

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies*, 29(3), 155–173. DOI: 10.2307/2295952.

Belderbos, R., Carree, M., Diesderen, B., Lokshin, B., Veugelers, R. (2004). Heterogeneity in R&D Cooperation Strategies. *International Journal of Industrial Organization*, 22(8–9), 1237–1263. DOI: 10.1016/j.ijindorg.2004.08.001.

Belderbos, R., Gilsing, V., Lokshin, B. (2012). Persistence of, and Interrelation between, Horizontal and Vertical Technology Alliances. *Journal of Management*, 38(6), 1812–1834. DOI: 10.1177/0149206310386962.

Cappellari, L., Jenkins, S. P. (2003). Multivariate Probit Regression Using Simulated Maximum Likelihood. *Stata Journal*, 3(3), 278–294.

Cohen, W. M., Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128–152. DOI: 10.2307/2393553.

Chaminade, C., Edquist, C. (2006). From Theory to Practice. The Use of the Systems of Innovation Approach in Innovation Policy, in Hage, J., De Meeus, M., eds., *Innovation, Learning and Institutions*. Oxford: Oxford University Press.

Colombo, M. (1995). Firm Size and Cooperation: the Determinants of Cooperative Agreements in Information Technology Industries. *International Journal of the Economics of Business*, 2(1), 3–29. DOI: 10.1080/758521094.

Dosi, G. (1988). Sources, Procedures, and Microeconomic Effects of Innovation. *Journal of Economic Literature*, 26(3), 1120–1171.

De Bondt, R. (1997). Spillovers and Innovative Activities. *International Journal of Industrial Organization*, 15(1), 1–28. DOI: 10.1016/S0167-7187(96)01023-5.

Duguet, E., Monjon, S. (2002). Creative Destruction and the Innovative Core: Is Innovation Persistent at the Firm Level? UCL. Discussion Paper 02–07.

Dyer, J. H., Singh, H. (1998). The Relational View: Cooperative Strategy and Sources of Inter-Organizational Competitive Advantage. *Academy of Management Review*, 23(4), 660–679.

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *Americal Journal of Sociology*, 91(3), 481–510. DOI: 10.1086/228311.

Gulati, R. (1998). Alliances and Networks. *Strategic Management Journal*, 19(4), 293–317. DOI: 10.1002/(SICI)1097-0266(199804)19:4%3C293::AID-SMJ982%3E3.0.CO;2-M.

Heckman, J. J. (1981a). Statistical Models for Discrete Panel Data, in Manski, C., McFadden, D., eds., *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge (MA): MIT Press, pp. 114–178.

Heckman, J. J. (1981b). The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, in Manski,

C., McFadden, D., eds., *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge (MA): MIT Press, pp. 179–195.

Hildrum, J. (2010). Inter-Organizational Learning in Drifting Environments—Experiences from a Multi-Firm Software Development Project. University of Oslo. TIK Working Papers on Innovation Studies with No. 20100122.

Mairesse, J., Mohnen, P. (2010). Using Innovation Surveys for Econometric Analysis, in Hall, B. H., Rosenberg, N., eds., *Handbook of the Economics of Innovation*, 2, pp. 1129–1155. DOI: 10.1016/S0169-7218(10)02010-1

Miotti, L., Sachwald, F. (2003). Co-Operative R&D: Why and with Whom? An Integrated Framework of Analysis. *Research Policy*, 32(8), 1481–1499.

Naylor, J. C., Smith, A. F. M. (1982). Applications of a Method for the Efficient Computation of Posterior Distributions. *Journal of the Royal Statistical Society,* Series C, 31(3), 214–225. DOI: 10.2307/2347995.

Lundvall, B. A., Johnson, B. (1994). The Learning Economy. *Journal of Industry Studies*, 1(2), 23–42. DOI: 10.1080/13662719400000002.

Lundvall, Borras (2005). Science, Technology, and Innovation Policy, in Fagerberg, J., Mowery, D. C., Nelson, R., eds., *Oxford Handbook of Innovation*. Oxford: Oxford University Press, pp. 599–631.

Nelson, R. R., Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. Cambridge (MA): Harvard University Press.

Noteboom, B. (1999). Innovation and Inter-Firm Linkages: New Implications for Policy. *Research Policy*, 28(8), 793–805. DOI: 10.1016/S0048-7333(99)00022-0.

OECD (1997). *Oslo Manual*. 2nd Ed., Paris: Organization for Economic Cooperation and Development.

OECD (2005). *Oslo Manual*. 3rd Ed., Paris: Organization for Economic Cooperation and Development.

Peters, B. (2009). Persistence of Innovation: Stylised Facts and Panel Data Evidence. *Journal of Technology Transfer*, 34(2), 226–243. DOI: 10.1007/s10961-007-9072-9.

Raymond, W., Mohnen, P., Palm, F., Van der Loeff, S. S. (2009). Innovative Sales, R&D and Total Innovation Expenditures: Panel Evidence on their Dynamics. UNU-MERIT. Working Paper No. 029.

Raymond, W., Mohnen, P., Palm, F., Van der Loeff, S. S. (2010). Persistence of Innovation in Dutch Manufacturing: Is it Spurious? *Review of Economics and Statistics*, 92(3), 495–504. DOI: 10.1162/REST_a_00004.

Rogers, M. (2004). Networks, Firm Size and Innovation. *Small Business Economics*, 22(2), 141–153. DOI: 10.1023/B:SBEJ.0000014451.99047.69.

Sachwald, F. (1998). Cooperative Agreements and the Theory of the Firm: Focusing on Barriers to Change. *Journal of Economic Behaviour and Organization*, 35(2), 203–225. DOI: 10.1016/S0167-2681(98)00054-7.

Soete, L., Verspagen, B., Weel, B. T. (2010). Systems of Innovation, in Hall, B., Rosenberg, N., eds., *Handbook of the Economics of Innovation*. Amsterdam: North Holland, pp. 1159–1180. DOI: 10.1016/S0169-7218(10)02011-3.

Srholec, M. (2009). Does Foreign Ownership Facilitate Cooperation on Innovation? Firm-Level Evidence from the Enlarged European Union. *European Journal of Development Research*, 21(1), 47–62. DOI: 10.1057/ejdr.2008.9.

Stata (2009). *Stata Longitudinal-data/Panel-data Reference Manual Release 11*. College Station (TX): Stata Press.

Stewart, M. (2007). The Interrelated Dynamics of Unemployment and Low-Wage Employment. *Journal of Applied Econometrics*, 22(3), 511–531. DOI: 10.1002/jae.922.

Sutton, J. (1991). *Sunk Costs and Market Structure*. Cambridge (MA): MIT Press.

Tether, B. (2002). Who Co-Operates for Innovation, and Why. An Empirical Analysis. *Research Policy*, 31(6), 947–967.

Van Leeuwen, G. (2002). Linking Innovation to Productivity Growth Using two Waves of the Community Innovation Survey. OECD Science. Technology and Industry Working Paper No. 8. OECD Publishing.

Veugelers, R. (1997). Internal R&D Expenditures and External Technology Sourcing. *Research Policy*, 26(3), 303–315. DOI: 10.1016/S0048-7333(97)00019-X.

Veugelers R., Cassiman, B. (2004). Foreign Subsidiaries as a Channel of International Technology Diffusion: Some Direct Firm Level Evidence from Belgium. *European Economic Review*, 48(2), 455–476. DOI: 10.1016/S0014-2921(02)00327-6.

Wooldridge, J. M. (2005). Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics*, 20(1), 39–54. DOI: 10.1002/jae.770.

Woolthuis, R. K., Lankhuizen, M., Gilsing, V. (2005). A System Failure Framework for Innovation Policy Design. *Technovation*, 25(6), 609–619. DOI: 10.1016/j.technovation.2003.11.002