Panel data or pseudo panels for longitudinal research? Cross-national comparisons using the example of firms' training spend.

Abstract

Design/Methodology/Approach

This paper applies a pseudo panel methodology to investigate the evolution of financial investments in training at the firm level over time. This approach enables a more meaningful exploration of inter-temporal changes in situations where longitudinal data does not exist.

Purpose

The evolution of firm level practices over time has always been a keen area of interest for management scholars in general and more specifically scholars of HRM. However, in comparison to other social scientists, particularly economists, the relative dearth of firm level panel data sets has restricted the methodological options for exploring inter-temporal changes. This paper utilises the repeating Cranet international data set, 10 countries over 18 years, and creates pseudo panels as a viable alternative to exploring firm level changes over time.

Findings

The analysis is framed within a varieties of capitalism lens and by adopting a more meaningful approach to examining changes over time it leads us to question some of the 'truisms' linked to firms expected behaviours within different national institutional frameworks. In this case, doubt is cast over the expectation that firms in coordinated market economies are likely to display a stronger financial commitment towards training their workforce. In addition, it also indicates that the

relationship between training spend and the business cycle may not be as strong as previously thought.

Research Limitations/Implications

As with any large-scale quantitative analysis, it would always benefits from a larger number of observations and/or a longer time period, in this instance access to annual data rather than 4 or 5 year intervals would have been helpful. In addition being able to use survey weights within the analysis would have been useful, but they are not available for this data set. However, it does at least offer a viable methodological approach for scholars in scenarios where longitudinal data is highly unlikely to be available over lengthy time periods.

Practical Implications

By adopting a different, and more appropriate, approach to analysing existing cross-sectional data over time this empirical research helps to achieve a deeper understanding of the complex issues that influence decision making at the firm level. Over time, this will enable policy makers to become better informed of the factors determining change over time.

Social Implications

At the firm level, in line with the practical implications above, this will enable decision makers to achieve a deeper understanding of the evolution of the external context in which they operate and the likely influence of that evolution within their own organisation. It will subsequently be able to inform more effective decision making and management of the human resource within those organisations.

PANEL DATA OR PSEUDO PANELS FOR LONGITUDINAL RESEARCH?

CROSS-NATIONAL COMPARISONS USING THE EXAMPLE OF FIRMS' TRAINING SPEND.

Introduction

One of the most important developments within quantitative empirical analysis of management policies and practices over the last three or four decades has been the increased use, and wider applications, of longitudinal data, enabling the more effective analysis of the dynamics of change within any population. This progress has to be tempered though by the usual criticisms that, although it includes repeated observations of the same respondents and allows change within the same sample to be explored, attrition rates make it very difficult to sustain a meaningful panel over any extended period. In addition, any analysis simply highlights behaviour and changes to that behaviour amongst the panel respondents, who may not necessarily continue to be representative of the overall population. In addition, the problem of persistent measurement error has been highlighted (Ashenfelter 1983). A counterview, which we present here, is that repeated cross-sections, regularly updated to ensure representative samples, may well paint a more realistic picture of changes over time.

Longitudinal panel data undoubtedly facilitates the exploration of different features. For example, in macroeconomics the relationship between growth of the economy and domestic consumption is a key determinant of the effectiveness of any expansionary/ contractionary policy intervention and the propensity to consume from any additional income is the key factor in this relationship. In the USA, once the Panel Study of Income Dynamics data set became available for empirical analysis, it was clear that previous estimates of the propensity to consume using repeated cross-sections were significantly different from those emanating from panel data. Kuznets (1962) highlighted that in estimating elasticities of the coefficients many groups of goods and commodities were significantly higher as measured by cross-sections than they were as measured by time series data, suggesting that the responsiveness of demand to changes in income were not as high as previously thought. As a consequence, a re-evaluation of the likely impact and effectiveness of various macroeconomic policies, plus an updating of the forecasting models, was required.

Although longitudinal data does support the exploration of different phenomena and the application of different methods, such data remain relatively uncommon, so the question of what to do in the absence of panel evidence still remains. However, since the seminal paper by Deaton (1985), the creation of a pseudo panel based on identifiable groups with fixed membership has offered a potential solution to this problem. The basis of the approach is that within the group individual observations can be replaced by group means as the basis for empirical analysis. A pseudo panel approach does lend itself to the analysis of firms' behaviour over time since the prevalence of longitudinal data sets is considerably rarer than with household surveys. The much higher attrition rates caused by company failures as well as mergers and acquisitions mean that the structure of any business panel is unable to survive for too long. For example, more recent waves of the Workplace Employment Relations Survey (WERS) data set include a subset of the data which are repeated collections from the same workplaces, i.e. a panel, however these are available for two consecutive waves only, presumably due to attrition, and as a result the analysis of inter-temporal changes over an extended period is impossible.

We take as an example of pseudo-panel data the repeating, internationally comparative, Cranet survey into human resource management (HRM). The Cranet data set fits our case as it is a repeated cross-sectional survey of organisations, representative at the national level by industry (based on the most relevant criterion for HRM, of employment) and by size (above 50 employees). The survey is repeated every 4-5 years. There are currently 6 waves of data collected between 1991 and 2013 available for analysis, hence they offer the potential for exploration of organisational behaviour and changes to that behaviour over a significantly longer time period than is available in most such surveys. The data set is also able to highlight the potential impact of unobserved group effects within that analysis. This paper seeks to make use of these data and takes as a specific a large scale inter-temporal and cross-country analysis of the financial commitment to training at the firm level using a pseudo panel approach. Training spend is chosen as the focus for the analysis since it is a key indicator of the extent to which the development of the human resource is pursued at the organisational level. In addition, there are likely to be distinct and predictable differences in the extent of the commitment to training across countries, as well as over time, and the following section will outline the theoretical underpinning for this.

Background

Over the last two decades, within the broad fields of international management and international HRM, it has become increasingly common to utilise the varieties of capitalism (VOC) literature as a means of rationalising and understanding cross-country differences in various types of behaviour at the firm level as well as within firms. Since the work of Whitley (1999), Hall and Soskice (2001) and Amable (2003) which developed and then refined the concepts of liberal market economies (LMEs) and coordinated market economies (CMEs) there have been numerous empirical studies exploring and highlighting the typical behaviours of firms within these differing national institutional frameworks (see Dilli et al 2018, Feldman 2019 and Hall 2018 as examples of some of the more recent applications). The basis of the distinction between the two types of economy is fundamentally the way that resources are allocated, with LMEs tending to be more reliant on market transactions and competition and CMEs, as the name suggests, have a greater tendency towards coordination between organisations and major stakeholders. The most important resultant distinction for the purpose of this study is that within the VOC economy types there is a tendency towards a very different relationship between the firm and the external market as well as a very different level of interdependence between the employer and its employees. In terms of the latter, LMEs are likely to be competitive for labour between themselves whereas in CMEs employees have stronger rights in their job and are more likely to stay with the employer for longer periods of time.

The upshot is that the various authorities expect that LMEs firms will be more willing to perceive and utilise the external labour market as a source of skills, i.e. for LMEs where particular skills are not available in sufficient quantities within a firm that firm is more likely to recruit externally to address that shortfall. By contrast, in CMEs firms are more likely to perceive their existing workforce as a source of skills and when faced with the same challenge of lack of skills a firm within a CME would be more likely to train and develop existing employees to address the problem. In archetypal CME firms, the external labour market is only typically used for recruitment into entry-level positions, to replace people who leave and in periods of business expansion. Therefore, for the purposes of the empirical analysis, this leads us to our main hypothesis;

H_1: Firms operating within CMEs are likely to invest proportionally more in training for their existing workforce than those in LMEs.

Data

The data used for our analyses come then from the repeating Cranet survey. Data is collected from the senior HRM person in each establishment, across both the public and private sectors, via a questionnaire that is completed by a representative sample of employers, stratified by employment numbers at the national level. For the purpose of this empirical analysis, data are used from the most recent waves and from 10 countries that were included in each such wave. The reasoning behind these two choices was firstly that the 10 countries were the ones which easily fitted into the various definitions of LMEs and CMEs with the countries being: Austria, Belgium, Denmark, Finland, Germany, Netherlands, Norway, Sweden Switzerland as the CME countries and the UK as the one representative of an LME. In the larger countries the data collected were a representative sample of the overall populations by employment, in the smaller countries these were full population surveys. The 5 waves give a sufficient time period (twenty years, corresponding to 1995, 1999, 2005, 2010, 2013) to enable us to identify any inter-temporal factors through the pseudo panel approach that may be missed within a pooled cross-section.

Method

When the analyses were performed using the pooled Cranet data, we relied on OLS estimation. The model can be written as follows:

$$Training\ spend_i = \beta_1 CME_i + \beta_2 X_i + \varepsilon_i$$

where the dependent variable is the annual training spend as a proportion of the total wage bill for the ith firm, CME is the key explanatory variable of interest which takes the value of 1 for coordinated market economies (Austria, Belgium, Denmark, Finland, Germany, Netherlands, Norway, Sweden Switzerland) and 0 for liberal market economies (UK). X_i includes a set of controls including union density, the extent of strategic HRM (coded in four categories as follows: *1)* not consulted, *2)* implementation, *3)* consultative, *4)* from the outset), size of the firm, industry (categorised as services and manufacturing), presence of a joint consultative committee (JCC), annual staff turnover and time (year dummies for 1995, 1999, 2005, 2010, 2015) and, ε_i is the error term. The reference category being a UK manufacturing company without a JCC in 1995 that does not consult HRM on strategic decisions.

In terms of the explanatory variables the CME dummy is included to facilitate the testing of the formal hypothesis and, hence, the remainder are included as controls. Size is included since there is a likelihood that larger and more complex organisations may well invest more in training since they have greater scope and capacity to develop people internally. Time dummies are added simply to pick up the likely influence of the business cycle and other trends upon economic activity. Strategic HRM, the proportion of employees who are trade union members and JCC variables are included since they all reflect the presence of an additional voice in the decision-making process that is likely to be supportive of investment in training. Finally, annual staff turnover is included, simply to control for the likelihood of higher initial induction costs that would be incurred with higher turnover rates. The descriptive statistics for all of these variables from the individual firm level data are recorded below in Table 1.

Following the OLS estimation with time dummies using the pooled data, the same model was replicated benefiting from the pseudo-panel approach. As explained earlier, pseudo-panels move the unit of analysis from individual units – in our case firms – to subgroups or cohorts of a population. One of the most important, and challenging, elements in constructing the pseudo-panels is to ensure that the sub-groups are defined by a set of characteristics that do not change or could be assumed to remain relatively stable over the time under consideration (Russell and Fraas, 2005; Verbeek, 2008; Meng et al., 2014). This ensures that while the firms within the subgroups might change over the years, the characteristics of the groups they belong to can be considered as stable. Given the difficulties associated with deciding the optimal sub-groups forming the pseudo panels, it is common to test different combinations (see, for example, Meng et al., 2014) – a practice which is also employed in this paper.

The decision on sub-groups involves a trade-off: while a greater number of sub-groups brings about increased heterogeneity of pseudo-panels, this also reduces the number of firms within each sub-group which may result in less precise estimates of sub-group means (Moffit, 1993; Verbeek and Vella, 2005). This becomes particularly challenging as firm-level data already suffer from smaller sample sizes when compared to national-level individual data sets. Nevertheless, in order to maximise precision and minimise measurement errors, it is necessary to have sufficiently large number of observations in sub-groups (cells).[1] Using simulation techniques and based on individual-level data, Verbeek and Nijman (1993) showed that cells must include about 100 individuals, though smaller cell sizes can be acceptable if the individuals within the cells are sufficiently homogenous. In

[1] When corresponding cells do not include the same individuals in two different periods, measurement errors could occur in pseudo panels. As noted in Gardes et al. (2005), if the first observation for cell 1 during the first period is an individual X, it will be paired with a similar individual Y observed during the second period. Therefore, measurement error arises between this observation of Y and the true values for X if X had been observed during the second period (Gardes et al., 2005: 243). However, this sort of measurement error is insignificant when cell sizes are large (Moffit, 1993; Verbeek and Nijman, 1993; Gardes et al., 2005) and, our cell size restrictions of minimum 30 firms should eliminate such concerns. Moreover, our unit of analysis are firms which may also reduce the extent of any measurement error caused by not having the same firms across the periods – it is likely that the group of firms with same characteristics will be less prone to measurement error in comparison to the group of individuals.

contrast, when sample size across the subgroups (cells) are too large, then there is a risk of loss of efficiency of the estimators.

The solution proposed for such challenges associated with creating pseudo-panels is to generate optimal groups where the loss of efficiency is reduced and the measurement error is negligible (Baltagi 1995; Gardes et al., 2005). This requires a thorough consideration when defining potential cohort/subgroups and ensuring that heterogeneity within them is minimized whilst heterogeneity between them is maximized (Cramer, 1964; Verbeek and Nijman, 1993). Once these empirical principles are followed and implemented to the data of interest carefully, consistent and efficient estimators can be obtained using pseudo-panels (Gardes et al., 2005).In light of the existing work, we tested the validity of several combinations to ensure that firms within the Cranet data are grouped into homogeneous and sufficiently large groups to achieve the most precise estimates.

Some examples of our initial attempts include creating pseudo panels with detailed industry categories[2] and the use of individual countries rather than the CME and LME sub-categories in addition to the standard candidates for grouping the firms (such as sector, union membership and so on). However, these combinations resulted in very small cell sizes and it was not possible to achieve precise estimates.[3]

After testing the validity of several combinations, the subgroups of our pseudo panel were created based on two types of market economies (CMEs and LMEs), five time periods (years 1995, 1999, 2005, 2010, 2015), four strategic HRM characteristics, two industries (services and manufacturing),

[2] Rather than aggregating the industries into services and manufacturing, we benefited from the more detailed industry categories available in the Cranet data including sub-divisions within the manufacture and service sectors (e.g., non-energy chemicals, metal manufacture, other manufacture, banking and finance, personal services, other services).

[3] To give an example, when individual countries, union membership, industry, strategic HRM, public/private sector and the presence of JCC were used in forming the sub-groups, 70.24% of the subgroups turned out to have sample size of 30 firms or less. On the contrary, when CME, industry, public/private sector and the presence of JCC were used in creating the subgroups, sample sizes within the groups were no longer an issue (more than 95% of the subgroups had more than 30 firms). Nevertheless, this resulted in a limited number of subgroups (39 subgroups with more than 30 firms) and less variation between them. Results from these attempts can be provided by the authors upon request.

two sectors (public and private) and presence of JCC. Although this would normally produce 320 sub-groups, given that some combinations lacked observations, we ended up with 236 sub-groups. Another caveat relates to the sample size for each sub-group. Whilst our intention was to exclude sub-groups with less than 100 observations as suggested by Verbeek and Nijman (1993) as optimal minimum; this would mean excluding a disproportionately large portion of the data. Nevertheless, as argued by the authors, smaller cell sizes could well be justified in cases where sub-groups are relatively more homogenous. In parallel, there are examples of pseudo-panels including smaller sub-group sizes (for instance, Propper, Rees and Green, 2001; Meng et al. 2014) or without any restrictions in sub-group sizes in case of repeated cross-sectional firm-level data (Brookes at al, 2017). Here, following the examples in the literature, we excluded those with less than 30 observations which is justifiable and pragmatic given the firm-level data.[4] This reduced the total number of subgroups to 61. This exclusion was necessary for a precise estimation which meets the required asymptotic properties for the pseudo-panel approach (see Verbeek, 2008 for a detailed discussion). Exclusion of subgroups with less than 30 observations resulted in losing a quarter of the total number of subgroups. We considered this an optimal minimum as excluding further sub-groups could have meant decreasing the heterogeneity of our pseudo-panel. Nevertheless, the results from the unrestricted panel (with no restrictions on cell size) will also be presented for completeness and comparison purposes

Once the pseudo-panels are created, the next step includes applying the standard panel data methods. We considered using the two most common methods adopted in analysing the panel data: fixed effects and random effects. Whilst random effects assume that unobserved individual effects (α_i) are not correlated with the independent variables, fixed effects allow for an arbitrary correlation between the two (Woolbridge, 2009). However, with fixed effect models, we cannot estimate the effects of time invariant variables whereas our key explanatory variables of interest (variables for

[4] .

CMEs and LMEs) are, by definition, time invariant. In fact, several other important control variables in our empirical specification are time-invariant. Acknowledging that explanatory variables being time invariant do not offer a full justification for choosing random effects over fixed effects, we performed a Hausman test to differentiate between the two models. Hausman test pointed the use of random effects as preferred model. [5]

The random effects model could be expressed as follows:

$$Training\ spend_{it} = \beta_1 CME_{it} + \beta_2 X_{it} + \alpha_i + \varepsilon_{it}$$

where, different from the previous OLS model with pooled data, the time dimension is introduced, and our dependent variable is now the training spent by the ith firm at time t. α is the unobserved time invariant individual effect.

Finally, in order to test for random effects, we applied Breusch-Pagan Lagrange multiplier (LM) test which helps to distinguish between the use of random effects or a simple OLS regression. The test indicated the use of random effects as an appropriate method.

Findings

The results from the pooled data and pseudo panels are presented in Table 2.

The first column of Table 2 shows the results based on the OLS estimation with pooled data. CMEs were found to spend more on training when compared to LMEs. A greater commitment to strategic HRM was linked to an increased amount spent on training. Firms in service industries were likely to spend more on training but the opposite was true for those in the public sector. The presence of a JCC showed a negative association. Although this may be counterintuitive, since it would be anticipated that staff side voices on JCCs are likely to be supportive of additional training

[5] Fixed Effects Model is still estimated for comparison purposes. The effects of time variant explanatory variables were similar to the ones observed using Random Effects Model (i.e., positive and statistically significant coefficient for union density and labour turnover). The results from the fixed effect model and the Hausman test are available upon request.

opportunities, it may well also be the case that such voices and businesses are also more skilled in securing training funded by external sources and the direct training cost to the firm is thus reduced. Also, although the coefficient was small, union density had a negative effect on the training spend. This too may appear to be counterintuitive since a trade union is always likely to be another voice promoting additional training, but it also needs to be remembered that unions are often successful in being able to increase wage levels for their members hence, if the latter is proportionally greater than the former, training spend as a percentage of the total wage bill is likely to fall and not rise.

When the analyses were replicated using a pseudo panel approach, results from our preferred pseudo panel data (column 2, Table 2) indicated that training spending of firms operating in CMEs were no longer statistically different than those in LMEs. While the effect of many other explanatory variables such as industry, sector, presence of JCC remained similar across the models, pseudo panel estimations showed a positive role of union density in training spent, contrary to the results based on the pooled data. Additionally, the coefficient on firm size became statistically significant and presented a negative effect on the training spent. Finally, annual staff turnover was shown to be positively linked to training spending; an effect which was not evident in the OLS model.

In relation to the formal hypothesis, clearly the pooled data estimates strongly support the view that CME based firms typically invest proportionally more in staff training, whilst once we adopt the pseudo panel approach the estimated results no longer support this. The explanation for this may well be that once group effects are introduced into the analysis the unobserved elements of that are more important than differences across economy types, but what those group effects may be are unobservable within these data.

The other variable of interest is union density and the transition of the estimated coefficient from negative and significant in the pooled cross-section to positive and significant in the restricted pseudo panel model. In the first instance it has already been highlighted that if unions raise wages by proportionally more than training spend then a negative coefficient would result. However, once we have introduced group effects and the coefficient becomes positive, this may well suggest that unions have a much greater influence on training spend at the industry level rather than at the firm level.

At first glance it could well be concluded that the changes that result from moving from the pooled cross-section to the pseudo panel may simply reflect the smaller sample size, and resultant fall in explanatory power, inherent within pseudo panel estimates. This may well be the case with the unrestricted model (no restriction on cell size), reported in column 3 of Table 2, since virtually all the explanatory variables lose significance in comparison to the initial estimates. However, with the restricted sample reported in column 2, where the sample size is even smaller but the subgroups are more robust, most of the explanatory variables retain their level of significance. We would therefore argue that this supports the view of the pseudo panel as a viable and potentially more revealing approach to inter-temporal analysis.

Finally, it would be worthwhile to comment upon the time dummies presented in each model. As seen in Table 2, the coefficients for the time dummies are all positive and statistically significant; they increase gradually up until 2008 and drop slightly in 2013. This strong time trend is not immediately obvious in pseudo-panel estimations; we see the coefficients are statistically significant only for 2003 and 2013. Nevertheless, the coefficients are similar in magnitude and the higher standard deviations observed in pseudo panel might be due to the smaller sample size.

These results suggest that estimations based on pooled cross-sectional data might actually represent an accurate reflection of what has been happening in the economy over and above the impact of explanatory variables included in our models. However, it is also possible that what we are seeing in the pseudo panel is a combination of decisions made at the firm level as well as changes at the industry level. This would be individual firms choosing to invest more (or less) in training or industries expanding (or contracting) in line with the business cycle and them having a greater (or lesser) training need. Closer inspection of the data reveals the latter to be stronger. Across the whole sample the industry categories metal and other manufacturing fall from 32% to only 21% of the total firms in the data set between 1995 and 2013. Whilst in the categories banking and finance, health, and other services, industries that display higher proportionate spending on training throughout, the presence has increased from 21% in 1995 to 39% by 2013. As a result this suggests that the upward trend in training spend is less to do with individual enterprises opting to invest to a greater extent in training and more to do with the evolution of the economy. With this evolution involving a contraction in manufacturing and growth in services, with some of the service sector industries having a consistently greater training requirement. An overall implication of this is that the need for additional training is likely to increase still further if this process continues and, whether this training emanates from the state, sector or individual firms, there will need to be an increased commitment to training from somewhere within the economy.

Conclusion

Overall, there are three key things to take away from this paper and the related analysis. First, in situations where there are repeated cross-sections at the firm level, pseudo panels do offer a viable approach to estimating relationships over extended periods, at the very least as a means of checking and confirming findings from pooled cross-sections. Therefore, we would recommend that future researchers pursue this course of action wherever possible. Second, it does possibly start to question some of the 'truisms' associated with the varieties of capitalism literature. For example, it would largely be taken as a given that comparable firms in CMEs are on average likely to show a greater commitment to training than their counterparts in LMEs. However the pseudo panel estimates do at least suggest that this might not necessarily be the case, hence some of these relationships that have been supported to date via pooled cross-sectional analysis may need to be revisited in order to further explore some of the previously unexplained dynamics underpinning the relationships as well as potentially revealing what some of the unobserved group effects might be. Finally, it does help deliver a clearer understanding of changes over time, with the pseudo panel estimates highlighting the relative importance of sectoral expansion/contraction to the commitment to spending on training across the economy. With this implying that training spend within individual firms may not be as sensitive to changes in the business cycle as might be expected and, in addition, this is common across both CMEs and LMEs. Both of these final points support the view that future research in these areas needs to focus more clearly upon the impact of separate changes at the sectoral level and the firm level to changes in behaviour and practices within individual firms.

References

Amable, B. 2003. The Diversity of Modern Capitalism. Oxford University Press, Oxford.

Ashenfelter, O., 1984. Macroeconomic analyses and microeconomic analyses of labor supply, Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 21(1), pp 117-156, January.

Baltagi, B. H. 1988. An Alternative Heteroscedastic Error Component Model Problem. Econometric Theory, 4, 349-350.

Brookes, M., Brewster, C. and Wood, G. 2017. Are MNCs norm entrepreneurs or followers? The changing relationship between host country institutions and MNC HRM practices. International Journal of Human Resource Management, 28(12), 1690-1711.

Cramer, J. S. 1964. Efficient Grouping, Regression, and Correlation in Engel Curve Analysis. Journal of American Statistical Association, 59, 233-250.

Deaton, A., 1985. Panel Data from Time Series of Cross-Sections. Journal of Econometrics, 30, pp 109-126.

Dilli, S., Elert, N. and Hermann, A.M., 2018. Varieties of Entrepreneurship: Exploring the Institutional Foundations of Different Entrepreneurship Through Varieties of Capitalism. Small Business Economics, 51(2), 293-320.

Feldman, M., 2019. Global Varieties of Capitalism. World Politics, 71(1), 162-196.

Gardes, F., Duncan, G.J., Gaubert, P., Gurgand, M. and Starzec, C. 2005. Panel and Pseudo-Panel Estimation of Cross-Sectional and Time Series Elasticities of Food Consumption. Journal of Business & Economic Statistics, 23(2). 242-253.

Hall, P.A., 2018. Varieties of Capitalism in Light of the Euro Crisis. Journal of European Public Policy, 25(1), 7-30.

Hall P.A. and Soskice D., 2001. Varieties of Capitalism: The Institutional Foundations of Comparative Advantage. Oxford University Press.

Jackson, G. and Petraki, A., 2010. Understanding Short-termism: the Role of Corporate Governance. Glasshouse Forum.

Kuznets, S., 1962. Quantitative aspects of the economic growth of nations: VII, The share and structure of consumption. Economic Development and Cultural Change, 10, pp 1-92.

Meng, Y., Brennana, A., Purshouseb, R., Hill-McManus, D., Angusa, C., Holmesa, J. and Meier, SP. (2014). Estimation of own and cross price elasticities of alcohol demand in the UK—A pseudo-panel approach using the Living Costs and Food Survey 2001–2009. Journal of Health Economics.

Moffitt, R., 1993. Identification and estimation of dynamic-models with a time-series of repeated cross-sections. Journal of Econometrics 59, 99–123.

Propper, C., H. Rees and K. Green (2001), The Demand for Private Medical Insurance in the UK: A Cohort Analysis, The Economic Journal, 111, C180–C200.

Russell, J. E. and Fraas, J. W. (2005). An application of panel regression to pseudo panel data. Multiple Linear Regression Viewpoints, 31, 1–15.

Verbeek, M., 2008. Pseudo-panels and repeated cross-sections. In: Matyas, L.,Sevestre, P. (Eds.), The Econometrics of Panel Data. Springer-Verlag, Berlin, pp.369–383.

Verbeek, M., and Nijman, T. 1993. Minimum MSE Estimation of a Regression Model With Fixed Effects From a Series of Cross-Sections. Journal of Econometrics, 59, 125-136.

Verbeek, M. and Vella, F., 2005. Estimating dynamic models from repeated cross-sections. Journal of Econometrics 127, 83–102.

Whitley, R., 1999 Divergent Capitalisms: The social structuring and change of business systems, Oxford: Oxford University Press.

Table 1 Descriptive Statistics

Variable	Mean
Training spend (proportion of the total wage bill)	3.14
	(.75)
CME	0.69
JCC	0.79
HRM Strategy	
Not consulted	0.08
Implementation	0.05
Consultative	0.34
From the outset	0.52
Manufacture	0.43
Service	0.57
Size (log)	6.54
	(.36)
Turnover	6.61
	(3.75)
Union Density	48.6
	(17.18)
Number of Obs	6419

Note: Standard errors for continuous variables are in parenthesis.

Table 2 Pooled cross-section and pseudo-panel estimations

	Pooled data	Cell>30	No restriction
CME	0.297***	-0.365	0.231
	(0.109)	(0.399)	(0.289)
JCC	-0.377***	-0.672**	-0.562**
	(0.110)	(0.275)	(0.261)
Size (log)	-0.044	-0.576*	0.134
	(0.036)	(0.322)	(0.201)
Union Density	-0.003**	0.035**	0.013
	(0.001)	(0.015)	(0.008)
HRM Strategy			
Implementation	0.270	-0.073	0.298
	(0.201)	(0.377)	(0.325)
Consultative	0.346**	0.587**	0.164
	(0.163)	(0.274)	(0.316)
From the outset	0.677***	0.665**	0.676**
	(0.149)	(0.283)	(0.313)
Public	-0.345***	-1.274***	-0.368
	(0.116)	(0.403)	(0.291)
Service sector	0.360***	0.778***	0.021
	(0.100)	(0.292)	(0.243)
Turnover	-0.003	0.099**	0.051
	(0.007)	(0.048)	(0.034)
Year Dummies			
1999	0.319***	-0.090	0.152
	(0.123)	(0.315)	(0.354)
2003	0.604***	0.640*	0.299
	(0.127)	(0.336)	(0.395)
2008	0.749***	0.698	1.822***
	(0.164)	(0.430)	(0.420)
2013	0.661***	0.685*	0.722*
	(0.154)	(0.398)	(0.424)
Constant	2.751***	4.422**	0.919
	(0.296)	(2.151)	(1.393)
Observations	6,419	61	236
Number of groups		23	57

Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Base category for the Strategic HRM is the "not consulted" category.