

Exploiting administrative data to investigate where those leaving jobs get re-employed

Dunne, John

Administrative Data Centre, Central Statistics Office (CSO)

Skehard Rd

Cork, Ireland

Email: John.Dunne@cso.ie

Abstract

The paper will cover experiences from the Job Churn Explorer project at CSO with a particular focus on a sectoral flow analysis of job separations - where do those leaving jobs get re-employed. The project adapts and develops the underlying methodology outlined to date to the situation in Ireland to provide a detailed insight into the dynamics of job churn and its components as Ireland entered the current recessionary period. The analysis datasets used are derived from linking the following three sources

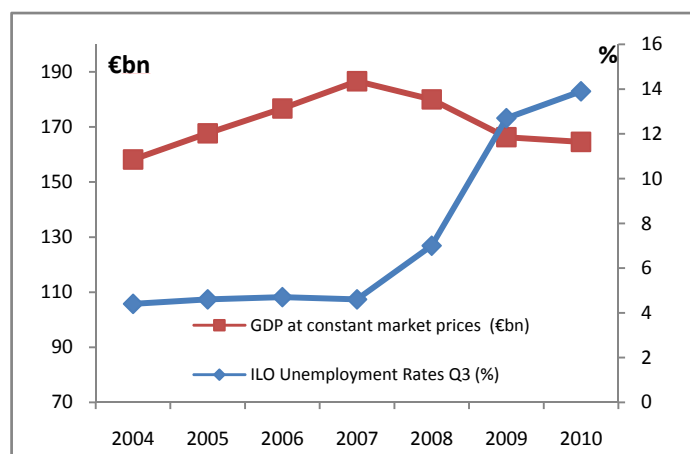
- Business register
- Employer tax returns
- Social Protection records

The comprehensiveness of the resulting analysis dataset, containing attributes on both workers and enterprises, provides for significant new opportunities to inform policy and decision making with respect to the labour market.

Focusing on job separations, (employer-employee relationships that existed in year t-1 but did not exist in year t) the paper will present the projects findings with respect to the sectoral flow of workers (the movement of workers within and between different sectors of the economy).

Introduction and Background

Ireland has, in recent times, experienced an unprecedented period of sustained growth followed by a sharp downturn at the end of 2007. This sharp downturn has had a significant effect on employment in Ireland with the unemployment rate rising from 4.6% in Q3 2007 to 13.9% in Q3 2010. GDP at constant market prices¹ fell nearly 12% from just over €186bn in 2007 to just under €165bn in 2010.



Jobs and any insights that can be obtained about the labour market that can inform decisions is, therefore, of significant value to government, business, workers and work seekers (Fox, 2009).

In particular, survey vehicles such as Labour Force Surveys are limited in their ability to track the flow of workers between jobs.

The work presented in this paper is informed by that in international literature and builds on the increasing relationship between CSO and

Government organisations with respect to exploiting administrative data for statistical purposes. In particular, this work allows for significant insights into job churn and its components in the Irish jobs markets, in other words the work provides information about those leaving, staying or taking new jobs and the firms in which these jobs are located.

Summary review of literature

There is a significant amount of literature available with respect to investigation of job churn and its respective components. Significant challenges are presented in much of the literature with respect to bringing the firm based components, job creation (JC) and job destruction (JD) together with the person based components, hirings (H) and separations (S) as typically they are derived from different sources. However, with the increasing recognition of the value of administrative data for statistical purposes, the use of employer-employee returns to tax authorities is of high value due to the 'single source' nature of the data when calculating and comparing the various components. Work has been identified in the US (Burgess, Lane, & Stevens, 2000), Finland (Ilmakunnas & Maliranta, 2001), Germany (Guertzgen, 2007) and Norway (Li, 2010) where employer-employee linked data sources has been able to facilitate a more comprehensive and in-depth insights into both job and worker components of job churn and how they interact with each other. The potential of such linked datasets is significant for obtaining insights into the movements of jobs and workers. These insights are of particular value to evaluating and informing policy analysts with respect to market dynamics for both jobs and workers.

Bassanini et al (2009) brings together from key papers and presents the underlying theory (including calculations and how they are derived) with respect to job churn in a clear manner with a view to bringing together results from many different studies and countries to undertake cross-country comparisons. The theory as presented has formed the basis of how the author has developed and calculated the various job churn components in this work.

¹ Chain linked and referenced to year 2008

Definitions and Methodology

The definitions and methodology used are adapted from those in (Bassanini & Marianna, 2009) to take account of methodology used in Eurostat-OECD Manual on Business Demography Statistics and shortcomings in the available data sources. The available data source does not have point in time measurements. Business Demography Statistics manual uses a methodology where year t is compared with year $t-1$.

The business unit of observation is that of an enterprise as defined in statistical legislation. Where administrative units have not been properly profiled into statistical units a one to one correspondence is assumed.

The primary variables for analysis at the business unit level are obtained by comparing data between two periods (calendar years) such that the following identity holds for each business unit

$$\Delta E_{it} = JC_{it} - JD_{it} = H_{it} - S_{it}$$

where E , JC , JD , H and S represent employment, job creation, job destruction, hirings and separations and Δ for differences between period $t-1$ and t .

Employment for the business unit in period t is estimated as the number of valid employment records with non zero reckonable pay² for that business unit in the period. This estimate does not factor in duration of employment or whether an employment is part-time or full-time in nature.

Job creation is measured as the difference in the number of employment records with non zero reckonable pay between two periods, t and $t-1$, if that difference is positive, zero otherwise and is assigned to period t .

Conversely, job destruction is measured as the difference in the number of employment records with non zero reckonable pay between two periods if that difference is negative, zero otherwise and is assigned to period t . In order for the identity to hold the jobs destroyed figures are assigned to period t even though technically the jobs were lost in period $t-1$.

Hirings for the business unit are calculated as the number of employment records assigned to an individual in period t for which a corresponding employment record for that individual did not exist in period $t-1$ with respect to the business unit.

Conversely, separations for the business unit are calculated as the number of employment records assigned to an individual in period $t-1$ for which a corresponding employment record for that individual did not exist in period t with respect to the business unit. Again, while technically the separations occur sometime in period $t-1$, for the identity to hold the estimated separations figure is assigned to period t .

Job stayers (JS) for the business unit are calculated as the number of employment records assigned to an individual in period $t-1$ for which a corresponding employment record exists for that individual in period t .

Job destruction figures for a group of business units is obtained by summing the figures for the business units in that group (i.e., for a group of business units classified to a specific sector). Job creation, hirings, job stayers and separations for a group of business units are also obtained in the same way.

Total job reallocation (REALJ) refers to the sum of job creation (JC) and job destruction (JD) for a group of business units. Excess job reallocation (EXCJ) for a group of business units is defined as the difference

² The primary difference between reckonable pay and gross pay is that reckonable pay excludes any payments to a pension schemes or permanent health insurance schemes recognised by the Irish Tax Authorities.

between total job reallocation (REALJ) and the absolute net change in total employment ($|JC - JD|$), for group j at period t.

$$EXCJ_{jt} = REALJ_{jt} - |\Delta E_{jt}| = JC_{jt} + JD_{jt} - |JC_{jt} - JD_{jt}|$$

Excess job reallocation provides a measure of the offsetting job creation and job destruction within a group of firms.

When aggregating over a group of business units with similar characteristics, generally speaking, job creation (JC) can be considered as the sum of employment growth from all expanding and new firms, while job destruction (JD) can be considered as the number of jobs lost from contracting or exiting firms. It should be noted that expanding and contracting business units are assigned these attributes based on volume or number of weeks work paid – therefore it is possible for contracting firms to have job creation and expanding firms to have job destruction (i.e., two employees each with 16 recorded weeks paid compared with one employee with 52 recorded weeks paid).

Worker reallocations are dealt with in a similar manner. Total worker reallocation (REALW) by summing hirings (H) and separations (S) over all members of a specified group, the group can be defined either by a group of firms or on a particular demographic characteristics (age, gender etc). Excess worker reallocation (REALW) for a group is defined as the difference between total worker reallocation (REALW) and the group's absolute net change in employment ($|H - S|$). So for group j at period t,

$$EXCW_{jt} = REALW_{jt} - |\Delta E_{jt}| = H_{jt} + S_{jt} - |H_{jt} - S_{jt}|$$

Excess worker reallocation provides a useful measure of the number of job matches over and above the minimum necessary to accommodate net employment growth; in other words, it reflects the reallocation of job matches (reshuffling of jobs and workers) within the same group (Bassanini & Marianna, 2009).

At the business unit level, churning flows (CH) is the difference between excess worker reallocation and excess job reallocation. Churning flows represent labour reallocation arising from firms churning workers through continuing jobs or employees quitting and being replaced on those jobs. So for group j in period t

$$CH_{jt} = EXCW_{jt} - EXCJ_{jt} = REALW_{jt} - REALJ_{jt} = H_{jt} + S_{jt} - JC_{jt} - JD_{jt}$$

All flow measures from period t-1 to period t are expressed as rates by dividing flow totals by relevant average employment figures in period t-1 and period t.

In adhering to recommendations in the literature, an average of the number of employments at year t and t-1 is used as the denominator in the calculation of rates with respect to reference period t.

Data sources

The datasets used in this analysis come from merging three separate sources as follows

- P35L data source from the Revenue Commissioners on employment records
- CRS Client Record System from the Department of Social Protection related to Personal Public Service Numbers (PPSN)
- CBR Central Business Register at CSO

The P35L is the primary source of data and contains a record for each registered employment i.e., employer/employee relationship, in the given year. The dataset contains an Employer Registration Number

(PREM number) that facilitates merging with the CBR to assign business based attributes and also contains the Personal Public Service Number and CRS to assign person based attributes. The P35L file also contains some records relating to Pension payments and these are excluded from the analysis. The P35L also contains information on number of weeks paid and reckonable pay (for tax purposes) for each employment record which can be used as indicators of job volume and value (and can be combined to give mean reckonable pay or an indicator of job quality). The P35L also contains the PPSN or person based public service identifier³. While this source can in general be considered exhaustive there are a small number of quality issues with respect to statistical purposes worth noting. In particular, the personal identifier when validated showed that for a small number of records (< 5%) an invalid number was recorded. However no significant pattern was identified to these invalid records. An interim decision was taken to work solely with records where the person number is identified to keep the methodology as simple as possible.

The CBR is the Business Register of enterprises maintained by CSO to support the compilation of statistics on business as laid down in EU statistical legislation. The business register became fully aligned to administrative data sources for reference year 2007. In general there is a one to one relationship between the enterprise as defined by the CBR ID and the employer registration number. However in a small number of cases an enterprise group may pass all of its employment through a single PREM number attached to a single enterprise. Another type of exception occurs where an enterprise can comprise of a number of legal units and hence have multiple PREM numbers. The CBR also does not have comprehensive coverage of all employment sectors. These difficulties arise due to the lack of a Unique Business Identifier across all public administration systems and also the lack of a standard methodology to profile enterprises in the Public Sector.

The CRS is a master register of all PPSNs assigned and contains information collected at registration on date of birth, sex and nationality as declared by the applicant. Nationality has only been collected since 2002. Any PPSNs assigned prior to this period are assumed a nationality of Irish for the purposes of creating the analysis datasets. This is done on the basis that prior to 2002 Ireland did not have the same influx of foreign nationals as it did after the enlargement of the EU to EU25. The PPSN came into being in 1998 and replaced the old RSI number used for tax and social welfare purposes. The purpose of the PPSN is to uniquely identify persons/customers when engaging or transacting with the state and is assigned when a person first interacts with the State. For those born in Ireland the PPSN is assigned shortly after birth (and is required to avail of child benefit). It is acknowledged that there are some quality issues with respect to PPSNs inherited from the old RSI number such as duplicate numbers, persons being assigned more than one RSI number or an identical RSI number (with a suffix of M or F) for husband and wife. However for statistical purposes these quality issues are not considered significant.

Two units of observation are available in the data for businesses, the first the Employer unit refers to the unique registration number of each employer while the second refers to the statistical definition of an enterprise in EU statistical legislation as applied in Ireland. The project uses the latter.

In summary, the CBR contributes the legal form and activity breakdown (NACE 1 and 2) attributes on the enterprise while the CRS contributes DOB, nationality and sex attributes on the person. This paper works with NACE Rev 2.

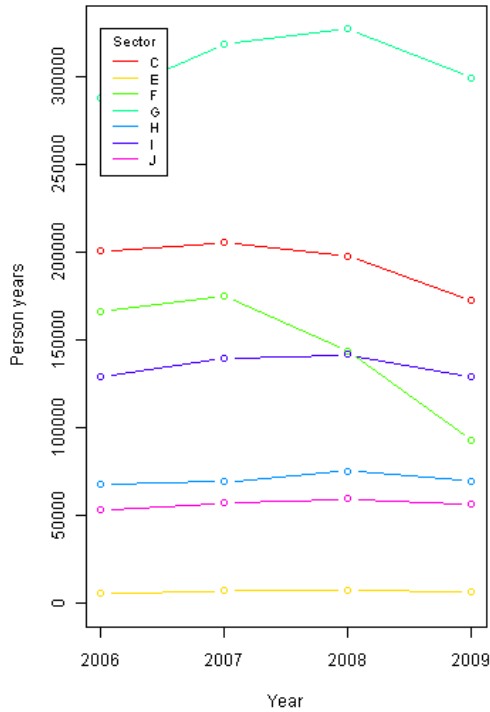
The P35 data source is available for reference years 2005 onwards. The CRS provides sufficient information on persons in the P35 files for reference year 2005 onwards.

³ In line with its data protocols, CSO replaces the official PPSN on analysis based datasets with a proxy for PPSN called the CSOPPSN. It is this proxy that is used to link person based data.

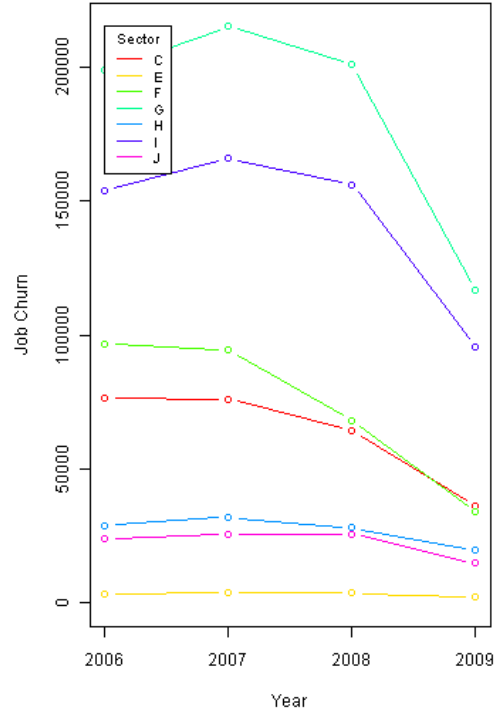
Investigating worker flows between sectors

- Sectoral flow of workers in context – job churn

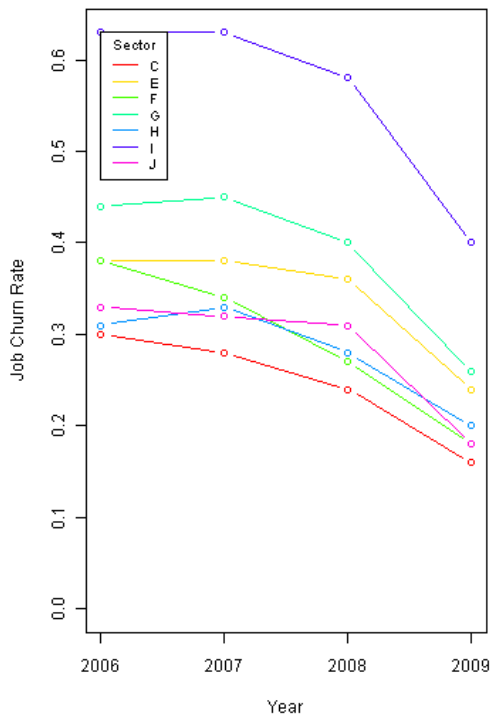
Volume of work (Person years) by Selected Sector



Job churn for Selected Sectors



Job churn rate for Selected Sectors



Looking at volume of work by sector (figure on top left), the largest and second largest sectors are the Wholesale and retail trade sector (G) and Manufacturing sector (C). The drop in employment since 2007 in the Construction sector (F) is also very apparent falling from approx 175,000 person years in 2007 to under 100,000 person years in 2009. For other significant sectors lesser declines in volume of employment are observed, commencing in 2007 for Manufacturing sector (C) and in 2008 for the Wholesale and retail trade sector (G) and the Accommodation and food sector (I). The figure at top right presents the absolute job churn figures or the movement of workers between firms above and beyond that required to satisfy the movement of jobs (job creation and job destruction) while the near left figure presents the job churn figures standardised as rates. The largest sector in terms of volume of work, Wholesale and retail trade (G) is also the sector providing the largest number of opportunities with respect to the movement of workers. The Accommodation and food sector (I) provides more opportunities for job movers/seekers than those in Manufacturing (C) despite being a significantly smaller sector in terms of volume of

work. In fact when at churn rates, Accommodation and Food (I) has a significantly higher rate than all other sectors with the Wholesale and retail trade (G) having the second highest churn rate. The sectoral analysis of churn rates is further evidence to the assumption that job churn is pro-cyclical with decreases in the job churn rate evident in all sectors.

- **Sectoral flow of workers in context – re-employability of separations**

Table 1 Analysis of separations⁴ by year and whether a new employment record was found

	No New employment		All		New employment	
	Number	%	Number	%	Number	%
Primary separations						
Business economy excluding activities of holding companies (B to N,-642)						
2006	156,572	32	483,788	100	327,216	68
2007	163,820	31	532,238	100	368,418	69
2008	205,006	35	586,201	100	381,195	65
2009	281,253	47	600,740	100	319,487	53
Business economy services excluding activities of holding companies (G to N,-642)						
2006	113,296	31	362,803	100	249,507	69
2007	112,657	29	395,016	100	282,359	71
2008	137,166	32	431,906	100	294,740	68
2009	187,274	43	438,509	100	251,235	57
Industry (B to E)						
2006	19,053	36	52,830	100	33,777	64
2007	22,152	38	57,998	100	35,846	62
2008	24,416	41	59,436	100	35,020	59
2009	35,027	51	68,436	100	33,409	49
Construction (F)						
2006	24,223	36	68,155	100	43,932	64
2007	29,011	37	79,224	100	50,213	63
2008	43,424	46	94,859	100	51,435	54
2009	58,952	63	93,795	100	34,843	37

When looking at the reference year 2006 in table 1 above, of the 484,000 separations in the business economy for the previous year, 68% or 327,000 were identified as being employed in some sector whether in the business economy (B to N, - 642)⁵ or not. For 2009, this ‘re-employability’ figure of 68% had fallen to 53%. The drop in re-employability however is significantly greater for those in the Construction sector (F) compared to other sectors with the re-employability figure falling from 64% in 2006 to 37% in 2009.

⁴ Note where a person is identified as having more than one separation, only the separation with the highest number of weeks paid is counted.

⁵ Note sectors A and O through U are excluded from the business economy analysis in this paper. The classification used is based on published Business Demography data for CSO, Ireland.

- Sectoral flow of separating workers

Table 2 Sectoral flow of separations in business economy finding re-employment

	B to N,-642		Sector C		Sector F		Sector G		Sector I		Sector N	
	Number	%	Number	%	Number	%	Number	%	Number	%	Number	%
Business economy excluding activities of holding companies (B to N,-642)												
2006	277,696	100	26,175	9	44,291	16	66,045	24	40,110	14	34,996	13
2007	315,185	100	28,565	9	47,820	15	74,205	24	44,253	14	41,540	13
2008	325,051	100	28,114	9	39,467	12	80,333	25	48,046	15	47,163	15
2009	270,112	100	25,792	10	22,024	8	78,360	29	43,788	16	37,207	14
Manufacturing (C)												
2006	26,461	100	7,793	29	4,125	16	5,385	20	1,726	7	2,655	10
2007	27,391	100	8,755	32	3,684	13	5,370	20	1,779	6	2,864	10
2008	27,492	100	8,290	30	2,613	10	6,021	22	1,879	7	2,879	10
2009	26,175	100	10,485	40	1,532	6	6,394	24	1,794	7	2,522	10
Construction (F)												
2006	40,398	100	3,005	7	24,765	61	3,036	8	1,392	3	4,142	10
2007	46,561	100	3,417	7	28,183	61	3,761	8	1,785	4	5,153	11
2008	46,995	100	3,997	9	24,866	53	5,037	11	2,397	5	5,633	12
2009	30,493	100	2,772	9	13,522	44	3,706	12	2,177	7	4,260	14
Wholesale and retail trade; repair of motor vehicles and motorcycles (G)												
2006	69,722	100	5,365	8	5,175	7	31,591	45	7,666	11	6,876	10
2007	74,930	100	5,400	7	4,689	6	35,564	47	7,716	10	7,449	10
2008	76,874	100	5,137	7	3,320	4	37,623	49	8,745	11	7,545	10
2009	72,824	100	4,860	7	1,751	2	42,571	58	7,479	10	6,320	9
Accommodation and food service activities (I)												
2006	50,730	100	2,784	5	2,704	5	11,087	22	22,407	44	4,994	10
2007	56,349	100	3,039	5	2,490	4	12,378	22	25,113	45	6,037	11
2008	55,831	100	2,899	5	1,825	3	12,683	23	26,066	47	5,607	10
2009	43,975	100	1,732	4	962	2	9,002	20	23,590	54	4,157	9
Administrative and support service activities (N)												
2006	40,526	100	3,873	10	4,102	10	7,251	18	4,138	10	10,654	26
2007	46,463	100	4,172	9	5,059	11	7,973	17	4,644	10	12,929	28
2008	54,437	100	4,250	8	3,966	7	9,428	17	5,382	10	18,826	35
2009	39,872	100	2,842	7	2,160	5	7,384	19	4,828	12	13,365	34

Table 2 describes the flow of workers between sectors in the business economy over time in absolute and percentage terms. Over the period 2006 – 2009 generally for each sector those changing jobs are more likely to take a new employment in the same sector, for example, the percentage finding re-employment that do not change sectors increases from 44% to 54% in the Accommodation and Food Sector (I), 29% to 40% in Manufacturing (C) and 45% to 58% in the Wholesale and retail trade sector (G). The exception is the Construction sector which shows a decrease from 61% to 44% of workers finding re-employment in the same sector over the period 2006 to 2009. As the re-employability of Construction workers in the construction

sector fell in 2008 and 2009 there was an increase in the proportionate flow of workers from the Construction sector into the Retail and Wholesale sector (G) up to 12% in 2009 from 8% in 2006, the Accommodation and food sector (I) up to 7% in 2009 from 3% in 2006 and the Administration and other activities sector (N) up to 14% in 2009 from 10% in 2006. In general the Wholesale and retail trade sector (G) is identified as the biggest recipient sector of cross sector flow of workers from other sectors. This is followed by Accommodation and food sector (I). The biggest cross sector flows happen between these two sectors, (G) and (I), with between 20% and 23% of re-employed workers from the Accommodation and food sector (I) finding re-employment in the Wholesale and retail trade sector (G) in any year and a reciprocal percentage flow of between 10% and 11% over the period 2006 and 2009. These are the two largest sectors in terms of job churn identified earlier.

Note the difference in the total number of primary separations in the business economy finding re-employment (eg Year 2006 = 327,216) in table 1 and the number of primary separations in the business economy finding re-employment (Eg year 2006 = 277,696) in table 2 is explained by those separations that find employment outside the business economy (i.e, in sectors A and O through U) when considering those separations that find re-employment.

Concluding remarks

This paper has presented summary statistics from the Job Churn project at CSO with a particular focus on those leaving jobs and on whether and where they go back into employment. Detailed statistical information from this project is available through CSO online databases at <http://www.cso.ie/px> in order to facilitate further exploration by researchers and policy analysts. This detail includes

- An economic activity breakdown (150 codes) as per Business Demography system across all datasets
- Job churn components described in this paper for each economic activity code and employment size class
- Age, sex and economic activity breakdown for Hirings, Separations and Job Stayers (measures also include Employment records, Value of reckonable pay, Volume of Work)
- Separations analysis (as per table 1) by economic activity, whether re-employed or not (measures also include mean weekly reckonable pay from separating employment)
- Sectoral flow analysis (as per table 2) of those separations finding re-employment categorised by whether mean weekly reckonable pay increased or not.

Potential uses of this information include (but are not constrained to)

- Informing on the demographic structure of employees by sector
- Providing input to labour costs analysis
- Providing information on gender pay gaps
- Identification of sectors providing job opportunities (quality of those jobs in terms of pay)
- Contributing to longer term evaluation of jobs policy

Further enhancements of the information provided in the online database could include

- Geographical flow of workers through looking at county of employer
- Breakdown of analysis into contracting and expanding firms
- Analysis by Country of Ultimate Controlling Interest (UCI) – foreign ownership
- Investigation of exporting enterprises

The work undertaken in this project is an illustration of the untapped potential hidden in administrative data systems across Public Authorities. The project also demonstrates the considerable added statistical value that is available through the linking of such data sources.

Bibliography

Bassanini, A., & Marianna, P. (2009). *Looking inside the perpetual motion machine: job and worker flows in OECD countries*. Retrieved from <http://www.oecd.org>.

Burgess, S., Lane, J., & Stevens, D. (2000). Job Flows, Worker Flows and Churning. *Journal of Labor Economics* , 18 (3).

Fox, R. (2009, June). *Job Opportunities in the Downturn*. Retrieved March 15, 2011, from <http://www.fas.ie/NR/ronlyres/9ABC5EE1-CF20-4AA5-ACA4-C5B81DD9FE5E/793/jobsdownturn96.pdf>

Guertzgen, N. (2007). *Job and Worker reallocation in German establishments: the role of employers' wage policies and labour market equilibriums*. Discussion paper, Centre for European Economic Research, Mannheim.

Ilmakunnas, P., & Maliranta, M. (2001). *The turnover of jobs and workers in a deep recession: evidence from the Finnish business sector*. Helsinki School of Economics and Business Administration; The Research Institute of the Finnish Economy. Helsinki: The Research Institute of the Finnish Economy.

Li, D. (2010). *Job reallocation and labour mobility among heterogeneous firms in Norway*. Working Paper, Ragnar Frisch Centre for Economic Research.