Life Course Centre Working Paper Series

ARC Centre of Excellence for Children and Families over the Life Course

Labour Protection and Informal Work: A Cross-National Analysis of European Countries, 2004-2012

Luz A. Flórez
Regional Unit of Economic Studies, Banco de la República de Colombia

Francisco Perales
Institute for Social Science Research, The University of Queensland

A more recent version of this paper was published as Flórez L and Perales F. (2016) Labour Protection and Informal Work: A Cross-National Analysis of European Countries, 2004-2012. International Labour Review, 155(4), 623-650

No. 2015-07
May 2015

Australian Government
Australian Research Council

NON-TECHNICAL SUMMARY

Informal employment is defined by the International Labour Organization as "economic activities by workers and economic units that are in law or in practice not covered or insufficiently covered by formal arrangements". In research, individuals who work without a written contract are usually considered to be informal workers.

Informal work has become a recognised problem in both developing and developed countries. This is because informal workers are not covered by the law, may lack access to unemployment insurance, pensions, or the health system, and are less productive. This is not only unfair to the individual worker, but it is also detrimental to the economy as a whole, as informal workers do not pay taxes. As a result, reducing informal work is a recognised policy objective.

During the last few decades and especially after the 2008 Global Financial Crisis (GFC) several European countries implemented employment reforms involving cuts to labour protection to increase labour market flexibility and reduce unemployment. While some authors argue that these measures worked in reducing unemployment rates, we posit that one of their 'side effects' was the rise in the prevalence of informal work.

Testing this is the goal of this paper. To do so, we analysed whether or not different indicators of labour protection were associated with the prevalence of informal employment in 20 European countries during the time period comprised between 2004 and 2012, i.e. before, during and after the GFC.

Key results indicate that increases in labour protection are associated with reductions in the prevalence of informal work. This finding has significant policy implications, and suggests that policies promoting labour market flexibility through reducing the generosity of unemployment benefit schemes and the regulation of employment protection can have the unintended consequence of increasing the size of the informal sector. This may offset (or partially offset) the benefits associated with labour market flexibility.

These findings can thus inform and contextualise current political and media debates about the pros and cons of implementing employment protection legislation in times of economic recession.

ABOUT THE AUTHORS

Luz A. Flórez is a Junior Researcher at the Regional Unit of Economic Studies at Banco de la República de Colombia. She has a background in economics, with research interests in macroeconomics, financial economics and labour markets in developing countries. Her research has appeared in Journals such as *Applied Economics, Cuadernos de Economía: Latin American Journal of Economics* and *El Trimestre Económico.* Her current research focuses on informal work in developing countries and the policies that governments can implement to reduce it. Email: lflorefl@banrep.gov.co

Francisco (Paco) Perales is Research Fellow in Family Dynamics at the Life Course Centre at the Institute for Social Science Research (The University of Queensland). His recent research has been on social disadvantage, gender inequalities, life course transitions, subjective wellbeing, and the use of household panel surveys within sociology. Recent publications include articles in gender attitudes and divisions of labour in *Social Forces* and *Social Science Research* and inequalities in wellbeing by sexual identity in *Social Indicators Research*. Email: f.perales@uq.edu.au

ACKNOWLEDGEMENTS: The authors would like to thank Arturo Martinez Jr, Tina Rampino and Blendi Kajsiu for helpful comments and suggestions. The Norwegian Social Science Data Services (NSD) is the data archive and distributor of the European Social Survey (ESS) data. Its Core Scientific Team (CST) and the producers bear no responsibility for the uses of the ESS data, or for interpretations or inferences based on these uses. The CST and the producers accept no liability for indirect, consequential or incidental damages or losses arising from use of the data collection, or from the unavailability of, or break in access to the service for whatever reason.

lifecoursecentre

(ARC Centre of Excellence for Children and Families over the Life Course)
Institute for Social Science Research, The University of Queensland
St Lucia, Qld 4072, Telephone: +61 7 334 67477
Email: lcc@uq.edu.au, Web: www.lifecoursecentre.org.au

Abstract

Informal work, defined as work that is undertaken without a formal contract, lowers productivity, reduces government tax revenue and hampers economic growth. Thus, reducing informal work is a recognised policy objective in developed and developing countries. Social and economic policy in Europe shifted markedly in recent years, particularly after the emergence of the Global Financial Crisis in 2008. Most reforms involved decreasing the generosity of unemployment benefit schemes and reducing the regulation of employment protection. We argue that, while these reforms may have contributed to reducing unemployment, they might also have had negative consequences, such as the increase of informal work. To test this, we use cross-national European Social Survey data for 2004-2012 augmented with external macro-level variables, and pooled and pseudo-panel regression models. Key findings indicate that labour protection is effective in reducing the prevalence of informal work amongst employees.

Keywords: labour protection; social security programs; informal employment; pseudo-panel estimation; cross-national analysis; Global Financial Crisis

1. Introduction and background

Informal employment, defined by the International Labour Organization as "*economic activities by workers and economic units that are in law or in practice not covered or insufficiently covered by formal arrangements*" (2002, p.53) is a recognised problem in both developing and developed countries. Informal workers are not covered by the law and lack access to unemployment insurance, pensions, or the health system. Informal workers are also less productive (Bernal 2009, Henley et al. 2009), which hampers economic growth, and do not pay taxes, which reduces government revenue and its subsequent ability to protect workers. As a result, reducing informal work has been a recognised policy objective in developed and developing countries for several decades.

During the last few decades and especially after the 2008 Global Financial Crisis (GFC) several European countries implemented employment reforms to increase labour market flexibility and reduce unemployment. Turrini et al. (2014) reviewed the breadth of policy reforms undertaken in European countries since the GFC, concluding that a majority of them involved decreasing labour protection. One way consisted of reducing the generosity of unemployment benefit schemes (Acemoglu and Shimer 1999a, 1999b; Cahuc 2014). This sort of reform has been more common in Eastern and Western Europe. Another way consisted of reducing the regulation of employment protection. These measures were most typically followed in countries with high unemployment rates, including Southern European countries such as Portugal or Italy. These courses of action are unsurprising given that labour protection restricts firing and hiring, increasing the costs incurred by employers to dismiss workers (Janiak and Wasmer 2014, Skedinger 2010). Altogether, there is evidence that these reforms did reduce unemployment and long-term unemployment rates (Martin 2014). However, by reducing labour protection, they might have had impacts on areas other than unemployment.

Empirical evidence on the macro-level effects of labour protection is mixed. Some studies find that labour protection has negative consequences. These include adverse effects on disadvantaged groups of workers (e.g. young people), increases in the number of workers in temporary contracts, reduced productivity, and lengthening of unemployment spells (Bajada and Schneider 2009, Skedinger 2010). Other studies report benefits of labour protection. For example, Janiak and Wasmer (2014) find a positive effect on capital-labour ratios through

skill protection. Similarly, Boeri et al. (2013) argue that tenure-based severance payments increase productivity by increasing investments in training, reducing inefficient firing, and promoting efficient labour allocation. There is also a literature highlighting the role of labour protection in reducing the negative outcomes of economic recession on the most vulnerable, on aspects that include mental health, mortality and suicide (Zivin et al. 2011, Suhrcke and Stuckler 2012).

Importantly, labour protection legislation can be used to protect workers against the decisions of employers, particularly in countries where employment law is sometimes not fully enforced (e.g. Southern European and Latin American countries):

> *"[A]n employer who does not comply with health and safety regulations in the workplace may fire workers who complain. The employer may have an interest to do so if he has monopsony power which allows him to replace those workers at low cost. Enacting a regulation which protects workers against such layoffs may improve efficiency."* (Cahuc 2014, p.12)

There is however a paucity of research devoted to systematically examining the effects of recent labour market reforms pertaining to labour protection on rates of informal employment across Europe, despite the importance of the subject matter and the number and depth of such reforms. Two studies stand as exceptions. Hazans (2011a) examined how labour market institutions are associated with growth in informal employment in Europe between 2004 and 2009. He finds that national rates of informal employment decrease with strict labour protection legislation and higher tax wedges on labour. Fialová (2010) examined the effect of labour institutions on the share of workers without a contract in Europe between 2003 and 2007. Her findings indicate that such share increased with the strictness of labour protection legislation. It must be emphasised, however, that Fialova (2010) defines informal work to include workers with temporary legal contracts.

Findings from these two prior studies are tentative, as they rely on aggregate country-level information. Additionally, they only relate to the period before the GFC, and so predate the more profound economic reforms that occurred in Europe thereafter. In this paper we contribute to the literature by analysing the impact of different indicators of labour protection

on the prevalence of informal employment in European countries between 2004 and 2012, using cross-national data from the European Social Survey (ESS) and pseudo-panel regression models. Key findings indicate that increases in labour protection are associated with reductions in the prevalence of informal work.

2. Theoretical model

We follow the theoretical model proposed by Flórez (2014). In this approach, time is continuous with $t \in [0, \infty)$ and workers die at rate $\mu > 0$, where μ also describes the inflow of new entrants in the labour market. For simplicity, all workers have the same subjective discount rate $\rho > 0$. There are two sectors within the economy: the formal sector and the informal sector. In the formal sector, workers earn the exogenous wage w. In the informal sector, workers earn the exogenous wage w_I. w depends on workers' productive capacities whereas w_I does not. We assume that w is higher than w_I.[1] While there are matching frictions in the formal sector, we assume that no such frictions exist in the informal sector. Following previous work in the field (see e.g. Maloney 1999, 2004; Albrecht et al. 2009), we also assume that the informal sector consists of unregulated self-employment. From this perspective, the decision to work in the informal sector is voluntary.[2]

In our model a representative worker has asset $A \geq 0$ and can be in one of two states: unemployed ($s = U$) and employed in the formal sector ($s = E$). Given their state, at every point in time the worker chooses consumption $c \geq 0$ where $u(c)$ describes the utility flow of consumption. The utility flow of consumption $u(c)$ is increasing, continuously differentiable and strictly concave with $u(0) = 0$. We also assume a perfect annuity market in which the worker enjoys the rate of return $r = \rho + \mu$ to savings and the worker's assets revert to the bank on death. The liquidity constraint $A \geq 0$ implies that banks do not lend to those with no assets.

[1] We do not consider cases when $w < w_I$ because, in this scenario, the decision to be an informal worker should not be affected by social security programs (Flórez 2014).

[2] Research in developing countries suggests that there is substantial heterogeneity within informal work, particularly between voluntary and involuntary informal employment (Fields 1990; Kucera and Roncolato 2008; Perry et al. 2007; Pagés and Stampini 2009). The latter is argued to be the most prevalent and detrimental.

When unemployed, the worker has three possible options. One option is to be non-participant, in which case he enjoys an additional utility flow $u_B > 0$ but his/her only income stream is asset income rA. A second option is to seek employment in the formal sector, in which case he enjoys an asset income rA plus a benefit flow $b \geq 0$ from the government. The jobseeker can also split his time between job search and employment in the informal sector and this split is unobserved by the government. For simplicity one of the following two possibilities is allowed: (i) searching full-time for employment and then receiving a formal job offer at rate λ (formal searchers), or (ii) taking casual employment which yields additional income w_I in the informal sector and only receiving a formal job offer at rate φλ with φ < 1 (informal searchers). A formal job offer implies that the worker becomes employed at wage w. The worker can quit without cost from employment and so becomes unemployed. Switching from unemployment to employment, however, requires search. Employed workers face job destruction shocks, which occur according to an exogenous Poisson process with parameter δ. The insurance program of the government is $B = \{b, S\}$, where b represents the social security payment,[3] and S is the lump sum received by workers when laid off.

The representative agent's Hamilton/Jacobi/Bellman equation that describes the value of being unemployed with assets $A \geq 0$ is given by:

$$rV_U(A) = max\begin{bmatrix} max_{c\geq 0}\left[u(c) + u_B + \frac{dV_u}{dA}[rA - c]\right], \\ max_{c\geq 0}\left[u(c) + \frac{dV_u}{dA}[rA + b + w_I - c] + \varphi\,\lambda\max[V_E(A) - V_U(A), 0]\right], \\ max_{c\geq 0}\left[u(c) + \frac{dV_u}{dA}[rA + b - c] + \lambda\max[V_E(A) - V_U(A), 0]\right] \end{bmatrix}$$

When $V_E(A) \geq V_U(A)$; i.e while it is (strictly) suboptimal to quit into unemployment, the

[3] In this theoretical model, b refers to *social security payments* rather than *unemployment insurance*. This is because governments cannot distinguish between individuals who (a) search for a formal job on a full-time basis and (b) search for a formal job on a part-time basis while working in the informal sector. Under these circumstances, social security payments provide a minimum level of resources during unemployment rather than insurance against lost earnings (Immervoll, 2012).

Hamilton/Jacobi/Bellman equation which describes the value of being employed with assets $A \geq 0$ changes to:

$$rV_E(A) = max_{c\geq 0}\left[u(c) + \frac{dV_E}{dA}[rA + w_I - c] + \delta \max[V_U(A+S) - V_E(A), 0]\right]$$

The last line in this equation describes the expected loss of receiving a 'shock destruction' with probability δ. When a worker loses his/her job, the level of assets increases by the severance payment S. These two Hamilton/Jacobi/Bellman equations are solved numerically in Flórez (2014). The solution suggests that labour protection should improve search incentives through 'entitlement effects',[4] protect workers from liquidity constrains, and ultimately reduce informal work. We test the last premise using suitable econometric methods described in the next section.

3. Econometric model

3.1 Pooled probit model

Using the above theoretical framework we test whether individuals' decisions to work in the formal or informal sector is affected by labour market policies including unemployment benefits, active labour policies and labour protection legislation. Such decision, conditional on current and future realizations of the exogenous variables, is given by:

$$n_{it} = 1(n_{it}^* > 0) = 1(x_{it}\beta + \mu_{it} > 0)(i = 1, \dots N; t = 1, \dots T) \quad (1)$$

where the i and t subscripts stand for individual and time respectively. n_{it} is an indicator function that is equal to 1 if the worker decides to be an informal worker and 0 otherwise, x_{it} is a vector of observed variables capturing socio-demographic and work-related

[4] *Entitlement effects* refer to the increase in unemployed workers' search intensity when unemployment benefits are about to exhaust (Mortensen 1977, Fredrikson and Holmlund 2001, Coles and Masters 2004, 2007).

characteristics, as well as social security programs implemented by the national government, and β is the corresponding vector of estimated model coefficients. $\mu_{it} = z_i + \varepsilon_{it}$ represents the combined effect of unobserved factors affecting the decision to become a formal or informal worker, and can be split into two components: time-invariant person-specific unobserved heterogeneity, z_i, and a standard stochastic error term, ε_{it}.

The above model can only be estimated in the presence of panel data, but only repeated cross-sectional data are available. In this scenario, two approaches are possible: pooled models and pseudo-panel models. Our first and most simple approach is to estimate a multivariate (pooled) probit regression model. This is given by:

$$n_i = 1(n_i^* > 0) = 1(x_i\beta + \mu_i > 0)(i = 1, \dots N) \tag{2}$$

This assumes that μ_i is uncorrelated with the x_i variables for the model to provide consistent and efficient estimates of the βs. Note that because we have data from different countries, the model includes dummy variables capturing country fixed-effects.

3.2 Pseudo-panel linear fixed-effect model

If the assumption of orthogonality between μ_i and x_i (a rather 'heroic' assumption) does not hold, then the pooled model gives inconsistent estimates of the βs. To relax this assumption, in the absence of suitable panel data we estimate pseudo-panel regression models (Deaton 1985). Additionally, pseudo-panel estimation has other important advantages over cross-sectional analysis: it enables combining data from multiple datasets into an integrated dataset, minimizes attrition bias, reduces bias due to individuals' response errors, and it minimizes the problem of 'ecological fallacy' by using both macro-level and individual-level data (Deaton 1985).

The pseudo-panel models take the form:

$$\bar{n}_{ct} = \bar{x}_{ct}\beta + \bar{u}_{ct}(c = 1, \dots C; t = 1, \dots T) \tag{3}$$

where $\bar{n}_{ct}$ is the average of n_{it} across individuals who belong to cohort c at time t, $\bar{x}_{ct}$ is the average of x_{it} across individuals in cohort c at time t, and $\bar{u}_{ct}$=$\bar{z}_{ct} + \bar{\varepsilon}_{ct}$ denotes unobserved heterogeneity.

If the number of individuals in each cohort is sufficiently large, then $\bar{z}_{ct}$ can be treated as a fixed parameter, assuming that variation over time can be ignored (Deaton 1985):

$$\bar{z}_{ct} = \bar{z}_c \tag{4}$$

Then, the pseudo-panel equation is given by:

$$\bar{n}_{ct} = \bar{x}_{ct}\beta + \bar{z}_c + \bar{\varepsilon}_{ct}(c = 1, \dots C; t = 1, \dots T) \tag{5}$$

where $\bar{z}_c$ represents time-invariant unobserved effects and $\bar{\varepsilon}_{ct}$ is the standard stochastic error term in regression. The within transformation is then applied to equation (5) to estimate the pseudo-panel fixed-effect model:

$$\bar{n}_{ct} - \bar{n}_c = (\bar{x}_{ct} - \bar{x}_c)\beta + (\bar{z}_c - \bar{z}_c) + (\bar{\varepsilon}_{ct} - \bar{\varepsilon}_c) \tag{6}$$

where $\bar{n}_c$, $\bar{x}_c$, $\bar{z}_c$, $\bar{\varepsilon}_c$ denote over-time cohort averages. Since $\bar{z}_c$ is by assumption time-invariant, it is averaged out of the equation:

$$\bar{n}_{ct} - \bar{n}_c = (\bar{x}_{ct} - \bar{x}_c)\beta + (\bar{\varepsilon}_{ct} - \bar{\varepsilon}_c) \tag{7}$$

Hence, under the assumptions posed above, the pseudo panel fixed-effect model eliminates the potentially biasing effects of time-constant unobserved heterogeneity, enabling more

robust estimation of the effect of labour protection on informal work. In our specific application, unobserved factors (such as unobserved skills and abilities, job preferences, personality traits and risk attitudes, or legal status of immigrants) might be correlated with both exposure to labour protection and informal work. It is for this reason that, unlike pooled cross-sectional models, pseudo-panel estimates will not suffer from omitted variable bias.

Note that the cohort mean of the individual-level outcome variable, i.e. whether the respondent works informally, is no longer a dichotomous variable but a proportion. Hence, the pseudo-panel fixed-effect regressions are linear models. This means that their results are not strictly comparable to those from the pooled probit models.

4. Data

We use multiple waves of data from the European Social Survey (ESS), specifically the available 2004, 2006, 2008, 2010, and 2012 instalments. The ESS is a multi-purpose large-scale repeated cross-sectional survey collected biannually through face-to-face interviews in over 30 European countries. Our interest is on temporal changes in informal employment, and so we restrict our analyses to countries for which the requisite information is available for at least 3 time points.[5] These include the following 20 countries: Belgium, Czech Republic, Greece, Switzerland, Germany, Denmark, Estonia, Spain, Finland, France, Hungary, Ireland, The Netherlands, Norway, Poland, Portugal, Sweden, Slovenia, Slovakia, and the United Kingdom. The total sample size in the analyses is *circa* 90,000 observations.

The ESS contains information on whether employees are working without a contract, derived from a question that reads: *"Do you have a formal contract of unlimited or limited duration?"*. We use this to construct our outcome variable: a dummy variable indicating whether a worker works in the informal sector (value 1), or the formal sector (value 0). Hence, as in Hazans (2011b), informal employees are those who have no verbal or written contract, whereas formal workers are those who have a verbal or written contract, irrespective of whether this is of limited or unlimited duration. Following research by the ILO (Hussmanns 2004), individuals who are self-employed (and by default do not have a contract)

[5] Croatia, Iceland, Latvia, Romania and Turkey were excluded from the analyses for this reason.

are catalogued as *formal* if they work in a professional occupations or employ more than five employees, and as *informal* workers otherwise. In the analyses, the subsamples of employees and self-employed workers are considered jointly as well as separately.

Table 1 shows the share of workers employed in the informal sector, by country. On average, about 19% of workers were informal workers across our sample of 20 European countries in 2004.[6] By 2012, this figure had fallen to about 16%. However, most of that change occurred in the 2004-2006 period, with little decreases and sporadic increases thereafter. Rates of informal employment are highly volatile across European countries. Over the sample period, these are generally higher amongst Southern European countries and lower in Northern European countries. In 2010 (the most recent year for which data are available across all countries), the countries with the highest informality rates were Greece (46%), Ireland (39%), Portugal (23%), United Kingdom (20%) and Spain (19%), while the countries with the lowest rates were Norway (9%), Estonia, France and Germany (10%). Both the share of informal employees and the share of self-employed workers decreased over time.

We augment the pooled ESS data by incorporating external measures of labour protection legislation by country and year. Following previous literature (Sapir 2006, Hazans 2011b, Fialová 2010) we use several different country-level measures of the intensity of social security programs derived from publicly-available OECD data. Social security programs provide financial support to low-income, disabled and unemployed individuals, and assistance to the unemployed to move into employment (Bajada and Schneider 2009). Reflecting this definition, empirical indicators of labour protection usually split the concept into two components: (i) protection of employed individuals and (ii) protection of unemployed individuals (Sapir 2006).

Table 1: Informal work as a proportion of total employment by country

	All workers					Employees					Self-employed workers				
	04	06	08	10	12	04	06	08	10	12	04	06	08	10	12

[6] Arguably, these figures are lower-bound estimates of the true levels of informal employment, as they do not take into account 'envelope wages', i.e. undeclared wages received by workers from their formal employers (Hazans 2011a).

Belgium	13	13	11	11	13	3	3	2	2	3	10	10	9	9	10
Czech Republic	13	-	13	15	14	3	-	2	3	2	10	-	11	13	12
Denmark	12	16	11	11	9	6	8	4	4	3	6	8	7	7	6
Estonia	9	9	11	11	10	2	5	4	3	3	7	4	7	7	7
Finland	11	11	11	13	12	1	1	1	1	2	11	10	10	12	10
France	-	10	11	10	-	-	4	3	2	-	-	6	8	8	-
Germany	12	13	12	10	10	2	2	1	2	1	10	11	11	8	9
Greece	53	-	52	46	-	25	-	23	16	-	28	-	29	30	-
Hungary	-	11	11	11	10	-	4	4	2	3	-	7	7	9	7
Ireland	43	41	39	39	36	29	27	22	23	20	14	14	16	16	16
Netherland	13	12	12	12	13	5	5	3	3	2	8	7	9	9	10
Norway	16	15	12	9	9	6	6	5	3	3	10	9	7	6	6
Poland	24	19	22	17	20	5	5	5	3	4	19	14	17	14	16
Portugal	20	23	24	23	25	4	9	9	8	10	16	14	15	15	15
Slovakia	14	13	12	14	15	3	2	1	1	3	11	11	11	13	12
Slovenia	13	12	15	11	13	6	5	7	3	2	7	7	8	8	11
Spain	19	21	20	19	19	5	6	5	3	5	14	15	15	16	14
Sweden	9	8	8	11	10	8	7	7	10	8	1	1	1	1	2
Switzerland	13	16	13	12	10	2	3	3	2	2	11	13	11	10	8
United Kingdom	24	22	22	20	21	13	12	11	9	8	11	10	11	11	13
All countries	19	16	17	15	15	7	5	5	4	4	12	11	12	11	11

Notes: ESS data (2004-2012).

We measure labour protection of employed individuals using the Employment Protection Legislation and Collective dismissal index (EPRC) devised by the OECD, which captures the degree to which regular workers are protected against individual and collective dismissal. This index incorporates different dimensions: notification, severance payments, and additional cost for collective dismissals, and can range from 0 (lowest protection) to 6 (highest protection) (see OECD 2013). This information is missing for Estonia and Slovenia for the years 2004 and 2006.

The second variable used to approximate the level of labour protection is focused on the protection of unemployed individuals, and operationalized as the amount of unemployment benefits to which workers are entitled net of taxes (i.e. the net replacement rate [NRR]). The NRR is given by the ratio of disposable income of unemployed individuals to their disposable income when they are at work (Salomäki and Munzi 1999). This variable can range from 0 to 1, where higher values denote higher subsidies to the unemployed net of taxes.

Additionally, we use a third variable to approximate the level of labour protection, namely national social expenditure as a proportion of GDP. This variable can also range from 0 to 1,

where higher values denote higher levels of labour protection (Hazans 2011b). [7] This complements the previous measures by capturing broader aspects of social protection, including social expenditure in active labor market programs, tax concessions, age pensions, disability schemes, the universality of the health system or housing protection, amongst others (Schneider and Enste 2000, Bajada and Schneider 2009).

Table 2 shows bivariate pairwise correlations between the measures of labour protection and the percentage of informal workers. Because these data vary only by country and year, country/year observations (n≈95) are used for the calculation of the Pearson correlation coefficients (r). As expected, the EPRC index (r=.08), degree of social expenditure (r=.21) and generosity of unemployment benefits (r=.48) are all negatively correlated with the percentage of informal workers, though the association is not statistically significant at the 10% level for the EPRC index. However, the 3 measures of labour protection are not very strongly correlated. *r* is .51 for employment protection and unemployment benefits, .37 for social expenditure and unemployment benefits, and just .17 for employment protection and social expenditure. The latter association is not statistically significant at the 10% level.[8]

Our multivariate analyses include the following individual-level control variables: self-employed status, gender, age, partnership status, parenthood status, whether retired, whether a full-time student, education, ethnicity, migrant status, occupational skill level, and employment sector. The models include also the following country-level control variables: country dummies (pooled model only), and per capita GDP growth. Additionally, all models control for study year using dummy variables.[9]

Table 2: Correlations between key variables

	EPRC	Unemployment benefits	Social expenditure
Unemployment benefits	.39***		

[7] For details on the multilevel data merged to ESS, see http://www.europeansocialsurvey.org/data/multilevel/

[8] We tried different methods to combine these indicators of labour protection, including principal component analysis and factor analysis. However, results were discouraging. For instance, the Cronbach Alpha statistic resulting from linearly combining the measures was only 0.34, way below the acceptable threshold of 0.7.

[9] Means and standard deviations for macro- and micro-level variables for all workers, employees and self-employed workers in the formal and informal sectors, and t-tests for sector differences are shown in Table A1 in the Appendix.

	(n=90)		
Social expenditure	.17	.32***	
	(n=90)	(n=94)	
% workers in the informal sector	-.08	-.33***	-.21**
	(n=90)	(n=94)	(n=94)

Notes: ESS data (2004-2012). Pearson correlation coefficients (sample size in parentheses). One observation per country/year. Significance levels: * = .1, ** = .05, *** = .01.

5. Empirical evidence

5.1. Pooled probit regression results

Table 3 presents the results of pooled probit models of the propensity to undertake informal work. The estimated effects are expressed as marginal effects, and so they give the predicted increase in the probability of a worker doing informal work associated with a 1-unit change in the explanatory variables.

Models (1) to (3) are estimated on the complete sample of workers, which comprises both employees and the self-employed. The key explanatory variable in Model (1) is the EPRC index. The estimated marginal effect on this variable (ME=−.041, p<.0.05) indicates that a 1-unit increase in the EPRC index is associated with a 4% reduction in the probability of a worker undertaking informal work, *ceteris paribus*. This effect is statistically significant at the 5% level. Therefore, employment protection, as measured by the EPRC index, decreases the size of the informal sector. Models (2) and (3) are analogous to Model (1) but substitute the EPRC index by other indicators of labour protection. Results suggest that neither the generosity of unemployment benefits (ME=−.035, p>.1) nor the degree of national social expenditure (ME=−.025, p>.1) have a statistically significant impact on the propensity of workers to undertake informal work.

Models (4) to (6) are estimated using the employee subsample, whereas models (7) to (9) are estimated using the self-employed subsample. The negative and statistically significant effect of the EPRC index on the propensity for workers to be part of the informal sector in model (1) emerges also amongst employees in model (4), and is more precisely estimated (ME=−.024, p<.01). However, the estimated effect is not statistically different from zero in the analogous model for the self-employed, model (7) (ME=.038, p>.1). This suggests that employment protection, as measured by the EPRC index, reduces the overall prevalence of

informal employment in a country by reducing the propensity for employees to take up employment in the informal economy, but does not affect the outcomes of the self-employed. Splitting the sample into employees and self-employed workers does not change the conclusions drawn about the effects on informal work of the generosity of unemployment benefits (models 5 and 8) or national social expenditure (models 6 and 9), for which the estimated marginal effects remain statistically insignificant.[10]

Altogether, results from pooled probit models suggest that, when measured as the EPRC index, employment protection is associated with a lower probability of employees working in the informal sector. For other labour protection indicators, we observe no statistically significant associations. However, the estimated coefficients on these models may suffer from omitted variable bias due to unobserved heterogeneity. In what follows, we present the results of pseudo-panel fixed-effect models which reduce such biases.

[10] The estimated coefficients on the control variables are for the most part consistent with theory and expectations. In most models, informal work is associated with per capita GDP growth, age, self-employment, being female, having children, being disabled, living in a rural area, being unpartnered, being retired, being a student, having low levels of education, and working in medium or low skilled occupations rather than low skilled occupations. There are also important differences in the prevalence of informal work across sectors of economic activity.

Table 3: Pooled probit model for working in the informal sector, marginal effects

	All workers			Employees			Self-employed		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*
Macro-level variables									
EPRC	-.041**			-.024***			.038		
	(-3.12)			(-4.79)			(0.54)		
Unemployment benefit		-.035			-.023			.123	
		(-0.70)			(-0.71)			(1.01)	
Social expenditure			.025			.008			.160
			(0.14)			(0.11)			(0.46)
Per capita GDP growth	.197*	.158*	.175*	.128**	.107**	.115**	-.186	-.122	-.131
	(2.58)	(2.21)	(2.24)	(3.17)	(2.72)	(2.69)	(-1.13)	(-0.94)	(-0.73)
Micro-level variables									
Self-employed	.380***	.379***	.379***						
	(39.50)	(39.63)	(39.74)						
Female	.018*	.018*	.018*	.010	.010	.010	.012	.012	.012
	(2.25)	(2.25)	(2.27)	(2.01)	(2.01)	(2.02)	(1.22)	(1.25)	(1.23)
Age/100	.050***	.050***	.050***	.025***	.025***	.025***	.062	.061	.061
	(2.44)	(2.41)	(2.43)	(2.46)	(2.43)	(2.45)	(1.29)	(1.29)	(1.28)
Partnered	-.012*	-.012*	-.012*	-.006**	-.006**	-.006**	-.016	-.016	-.016
	(-2.21)	(-2.24)	(-2.24)	(-3.10)	(-3.14)	(-3.14)	(-1.03)	(-1.04)	(-1.03)
With children	.006	.005	.006	.005***	.005***	.005***	-.012	-.012	-.012
	(1.86)	(1.83)	(1.82)	(3.53)	(3.43)	(3.42)	(-1.12)	(-1.11)	(-1.10)
Disabled	.048*	.048*	.048*	.031*	.031*	.031*	-.004	-.004	-.003
	(2.05)	(2.05)	(2.05)	(2.51)	(2.52)	(2.52)	(-0.14)	(-0.12)	(-0.11)
Lives in a rural area	.032***	.032***	.032***	.019***	.019***	.019***	.010	.010	.011
	(7.11)	(7.12)	(7.14)	(5.32)	(5.33)	(5.35)	(0.67)	(0.70)	(0.69)
Student	.063***	.064***	.064***	.033***	.034***	.034***	.006	.005	.005
	(6.83)	(6.86)	(6.86)	(6.83)	(6.85)	(6.84)	(0.44)	(0.41)	(0.41)
Retired	.047***	.048***	.048***	.031***	.030***	.031***	-.002	-.001	-.001
	(3.91)	(3.91)	(3.91)	(4.62)	(4.56)	(4.57)	(-0.10)	(-0.07)	(-0.06)
Education									
Tertiary (ref.)									
Primary	.044***	.044***	.044***	.020***	.020***	.020***	.094***	.094***	.095***
	(7.94)	(7.94)	(7.87)	(8.97)	(8.97)	(9.02)	(7.04)	(7.22)	(7.55)
Secondary	.019***	.019***	.019***	.006*	.006*	.006*	.076***	.076***	.076***

	(3.48)	(3.51)	(3.52)	(2.58)	(2.62)	(2.62)	(4.71)	(4.75)	(4.74)
Occupation skill level									
Low (ref.)									
Medium	-.036***	-.036***	-.036***	-.021***	-.022***	-.022***	-.006	-.006	-.006
	(-8.69)	(-8.67)	(-8.70)	(-11.32)	(-11.28)	(-11.36)	(-0.53)	(-0.55)	(-0.55)
High	-.096***	-.096***	-.096***	-.013**	-.013**	-.013**	-.288***	-.288***	-.287***
	(-4.71)	(-4.72)	(-4.72)	(-2.99)	(-2.99)	(-2.99)	(-14.64)	(-14.47)	(-14.46)
Local minority	.006	.005	.005	.003	.003	.003	.015	.015	.015
	(0.62)	(0.57)	(0.58)	(0.67)	(0.63)	(0.63)	(0.56)	(0.56)	(0.56)
Immigrant	.031*	.031*	.031*	.019*	.019*	.019*	.015	.015	.015
	(2.31)	(2.31)	(0.021)	(2.28)	(2.27)	(2.28)	(1.61)	(1.57)	(1.51)
Economic sector									
Manufacture (ref.)									
Agriculture	0.099***	.099***	.099***	.042***	.042***	.042***	.107***	.107***	.107***
	(9.26)	(9.29)	(9.28)	(9.09)	(9.05)	(9.07)	(4.08)	(4.10)	(4.11)
Construction	.045***	.045***	.045***	.021***	.021***	.021***	.073***	.073***	.073***
	(5.12)	(5.15)	(5.17)	(5.32)	(5.34)	(5.36)	(3.34)	(3.36)	(3.37)
Transport & trade	.013*	.012*	.013*	.004	.004*	.004*	.038*	.038**	.039**
	(2.53)	(2.58)	(2.60)	(1.99)	(2.03)	(2.04)	(2.41)	(2.44)	(2.46)
Public admin.	-.011	-.011	-.011	-.008	-.008	-.008	-.004	-.004	-.003
	(-0.87)	(-0.87)	(-0.87)	(-1.41)	(-1.42)	(-1.42)	(-0.22)	(-0.21)	(-0.19)
Survey year									
2004 (ref.)									
2006	-.005	-.003	-.004	-.003	-.002	-.002	.008	.005	.007
	(-0.66)	(-0.47)	(-0.56)	(-0.90)	(-0.63)	(-0.79)	(0.38)	(0.24)	(0.38)
2008	.002	.001	.002	.001	.000	.001	.015	.018	.0153
	(0.39)	(0.22)	(0.33)	(0.25)	(0.05)	(0.13)	(0.63)	(0.73)	(0.64)
2010	-.016**	-.015*	-.016***	-.011***	-.010**	-.011	.007	.005	.003
	(-2.71)	(-2.35)	(-3.25)	(-3.55)	(-2.94)	(-3.27)	(0.37)	(0.30)	(0.20)
2012	-.015*	-.010	-.011*	-.001	-.004	-.005	-.019	-.023*	-.024*
	(-2.08)	(-1.55)	(-1.88)	(-1.84)	(-1.07)	(-1.30)	(-1.41)	(-2.10)	(-2.07)
N (observations)	*87,130*	*90,048*	*90,048*	*74,706*	*77,388*	*77,388*	*12,424*	*12,660*	*12,660*
Log likelihood	*-17,805*	*-17,860*	*-17,860*	*-12,916*	*-12,961*	*-12,961*	*-3,828*	*-3,835*	*-3,836*

Notes: ESS data (2004-2012). Significance levels: * = .1, ** = .05, *** = .01. Control variables include country dummies. t-statistics in parentheses.

5.2. Pseudo-panel estimation

5.2.1. Constructing pseudo-panels

We use the repeated-cross sections of the ESS to build a pseudo-panel dataset following the approached detailed in Deaton (1985). This involves defining cohorts using time-constant individual characteristics, and dividing the population into different cohorts with the characteristics of a representative cohort member. Longitudinal analysis is then undertaken using the cohort as the units of analysis.[11]

We construct cohorts based on survey instalment (n=5), country (n=20) and workers' date of birth (n=6: 1946-52, 1953-59, 1960-66, 1967-73, 1974-80, 1981-87). This should result in 20*5*6=600 observations, from 20*6=120 cohorts, observed 5 times each. However, our pseudo-panels consist of fewer than 600 observations. This is because some countries did not participate in some ESS instalments,[12] or had missing data on labour protection variables, and two observations of the oldest cohort were excluded because they were based on fewer than 30 individuals (see Deaton 1985). The same process is repeated for the subsamples of employees and self-employed workers (although we do not restrict the cohort sizes to be over 30 individuals for these).

Table 4 shows descriptive statistics on the number of individuals used to construct the 562 observations from our 120 cohorts. For cohorts comprising all workers (i.e. employees as well as self-employed workers) the average number of individuals is 142.4. When considering only employees, the mean cohort size is still large, of 122.4. These numbers are well over the optimal cohort size of 100 individuals recommended by Verbeek and Nijman (1993). For the self-employed, however, the mean cohort size, 22, is very small. Hence, results using this subsample are to be interpreted with care.

Table 4: Cohort size

[11] There is an inherent trade-off between the number and size of cohorts (Deaton 1985, McKenzie 2004). Using a small number of cohorts maximizes cohort size, but may lead to inefficient estimation due to larger within-cohort heterogeneity. Using a large number of cohorts reduces cohort size, but may lead to biased estimation.

[12] Data are missing for Greece (2006, 2012), France (2004, 2012), Czech Republic (2006) and Hungary (2004).

Date of birth	Observations (cohort)	Mean	SD	Min	Max
All workers	562 (120)	142.4	51.4	31	356
Employees	562 (120)	122.4	44.3	8	294
Self-employed	555 (120)	20.2	13.0	1	96

Notes: ESS data (2004-2012).

5.2.1. Pseudo-panel fixed-effect regression results

Table 5 shows the results of pseudo-panel fixed-effect linear regression models of the propensity to work in the informal sector.[13] As explained before, results from these pseudo-panel fixed-effect models are superior to results obtained from pooled probit models because they account for time-invariant unobserved effects which, if correlated with the explanatory and outcome variables, would bias the estimated parameters. These are linear models, and so the estimated coefficients give the expected change in the proportion of workers in the informal sector within a cohort at a specific point in time associated with a within-cohort 1-unit over-time increase in the explanatory variables. Since the outcome variable is in fact a proportion, the coefficients approximate the within-cohort change in the probability that a worker is employed in the informal sector.

Models (10) to (12) are estimated on the complete sample of workers, which comprises both employees and self-employed workers. Unlike for the pooled probit models, all three labour protection indicators are significantly related to the probability of individuals working informally. The coefficient on the EPRC index (B=−.040, p<.1) indicates that a 1-unit increase in the EPRC index is associated with a decrease of 4 percentage points in the percentage of workers in the cohort employed in the informal sector, all else being equal. The coefficient on unemployment benefits (B=−.149, p<.01) indicates that a 1 percentage-point increase in the NRR is associated with a decrease of 1.5 percentage points in the percentage of workers working informally, *ceteris paribus*. Similarly, the coefficient on the

[13] The models are estimated using Stata routine `xtreg` with the `fe` option. Some propose that, when cohort sizes differ substantially, better estimation can be achieved using weights given by the square root of the cohort size (Devereux 2006, Ziegelhofer 2014). We tested the robustness of our results to this situation using Stata user-written subroutine `ppreg` (Lonkshin et al. 2008), which corrects for measurement error in the observed cohort means. The results are similar and are available upon request. Since `ppreg` cannot handle time interactions, we use the `xtreg` estimates as our main results.

degree of social expenditure (B=−.374, p<.01) indicates that a 1 percentage-point increase in social expenditure is associated with a 3.7 percentage point decrease in the percentage of workers employed in the informal sector, *ceteris paribus*.

Models (13) to (15) are estimated using only the employee subsample, whereas models (16) to (18) are estimated using only subsample of self-employed workers. Amongst employees, the EPRC index in Model 13 (B=−.042, p<.01), the generosity of unemployment benefits in Model 14 (B=−.184, p<.01) and the degree of social expenditure in Model 15 (B=−.493, p<.01) are all negatively and statistically significantly associated with participation in informal work. Once again, there are no statistically significant associations amongst the self-employed.

Overall, results from these preferred pseudo-panel specifications point towards similar conclusions as those drawn from pooled models: labour protection is associated with a decrease in the propensity for employees to undertake informal work. The results are also clearer, as all three measures of labour protection legislation have statistically significant impacts on the outcome. Since these pseudo-panel models account for cohort-specific unobserved effects, their results are more robust than those of the analogous pooled probit models. Thus, we take these as our preferred set of results.

Table 5: Pseudo-panel linear fixed-effect model for working in the informal sector models, model coefficients

	All workers			Employees			Self-employed		
	(10)	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*	*(16)*	*(17)*	*(18)*
Key macro-level variables									
EPRC	-0.040***			-0.042***			-0.031		
	(-3.19)			(-2.98)			(-0.64)		
Unemployment benefit		-0.149***			-0.184***			0.185	
		(-4.13)			(-4.78)			(1.32)	
Social expenditure			-0.374***			-0.493***			0.405
			(-2.71)			(-3.35)			(0.76)
N (observations)	*538*	*562*	*562*	*538*	*562*	*562*	*531*	*555*	*555*
N (cohorts)	*120*	*120*	*120*	*120*	*120*	*120*	*120*	*120*	*120*

Notes: ESS data (2004-2012). Significance levels: * = .1, ** = .05, *** = .01. Control variables as in Table 3 (except for country dummies). t-statistics in parentheses.

5.2.2. Changes in the effect of labour protection over time

We now examine whether and how the effects of labour protection policies on informal work have evolved over the 2004-2012 observation period. This is enlightening as to the effects of the GFC, which started in *circa* 2008. To do so, we estimate pseudo-panel fixed-effect regression models analogous to those in Table 5 which add interactions between survey year and labour protection measures. Selected results are presented in Models (19) to (21) in Table 6. Several interaction effects are statistically significant. These are more easily interpreted by looking at the margins reported in Table 7.

The pattern of results differs markedly across the different indicators of labour protection. The EPRC measure has a negative and statistically significant effect on the prevalence of informal work in years 2004 and 2006. Such effect becomes smaller in magnitude and statistically insignificant in years 2008-2012, coinciding with the emergence and establishment of the GFC. The generosity of unemployment benefits however has a comparable, statistically significant, negative effect on informal work over the complete 2004-2012 time period – though this is noticeably stronger in 2004. A similar pattern of results emerges for the social expenditure variable, which remains negatively and significantly associated with informal work through the observation window.

Overall, results from this exercise suggest that generous unemployment benefits and high levels of social expenditure reduced the prevalence of informal work before and during the GFC, whereas the importance of employment protection (as measured by the EPRC) eroded over time with the emergence of the GFC.[14]

[14] Identifying the specific effect of the GFC as an exogenous shock is beyond the scope of this paper and is complicated. For instance, the GFC hit different European countries at different times, which makes it difficult to put a starting date on it. Additionally, the GFC hit Europe after the United States, which means that complex anticipation effects may be at play. Hence, our results remain only tentative of these relationships. Further research on the *causal* effects of the GFC is warranted.

Table 6: Pseudo-panel linear fixed-effect model for working in the informal sector models with year*labour protection interactions (all workers), model coefficients

	(19)	*(20)*	*(21)*
Main effects			
EPRC	-0.043***		
	(-3.30)		
Unemployment benefit		-0.194***	
		(-5.13)	
Social expenditure			-0.388***
			(-2.64)
Survey year			
2004 (ref.)			
2006	-0.066***	-0.077**	-0.033
	(-2.83)	(-3.60)	(-1.37)
2008	-0.082***	-0.068**	-0.030
	(-2.74)	(-2.67)	(-0.85)
2010	-0.121***	-0.094**	-0.049
	(-3.46)	(-2.98)	(-1.28)
2012	-0.134***	-0.097**	-0.069
	(-3.21)	(-2.48)	(-1.60)
Interaction effects			
2006 * EPRC	0.020**		
	(2.44)		
2008 * EPRC	0.025***		
	(2.99)		
2010 * EPRC	0.033***		
	(3.65)		
2012 * EPRC	0.040***		
	(4.05)		
2006 * Unemployment benefits		0.096***	
		(3.51)	
2008 * Unemployment benefits		0.062**	
		(2.47)	
2010 * Unemployment benefits		0.076***	
		(2.85)	
2012 * Unemployment benefits		0.071***	
		(2.70)	
2006 * Social expenditure			0.040
			(0.39)
2008 * Social expenditure			-0.013
			(-0.12)
2010 * Social expenditure			-0.020
			(-0.18)
2012 * Social expenditure			0.030
			(0.28)
N (observations)	*538*	*562*	*562*
N (cohorts)	*120*	*120*	*120*

Notes: ESS data (2004-2012). Significance levels: * = .1, ** = .05, *** = .01. Control variables as in Table 3 (except for country dummies). t-statistics in parentheses.

Table 7: Impact of labour protection on informal work over time (all workers), margins

Year	EPRC	Unemployment benefits	Social expenditure
2004	-.043***	-.194***	-.388***
	(-3.30)	(-5.13)	(-2.64)
2006	-.023*	-.097**	-.349**
	(-1.75)	(-2.38)	(-2.19)
2008	-.018	-.132***	-.401**
	(-1.21)	(-3.33)	(-2.45)
2010	-.001	-.117***	-.408**
	(-.64)	(-2.93)	(-2.45)
2012	-.000	-.122***	-.358**
	(-.16)	(-3.10)	(-2.37)

Notes: Based on models 19, 10 & 21 in Table 6. ESS data (2004-2012). Significance levels: * = .1, ** = .05, *** = .01. t-statistics in parentheses.

5.2.3. Regional heterogeneity in the effects of labour protection

Finally, we explore regional differences in how different labour protection measures affect rates of informal work. To do so, we first construct dummy variables dividing countries into four geographical regions: Eastern Europe (Estonia, Hungary, Poland, Slovenia, Czech Republic, Slovakia), Western Europe (Belgium, Switzerland, Germany, France, United Kingdom, Ireland and The Netherlands), Northern Europe (Denmark, Finland, Norway and Sweden) and Southern Europe (Greece, Spain and Portugal). We then interact these dummy variables with the measures of labour protection, and add the latter to models analogous to those presented in Table 5. The new models, models (22) to (24), are shown in Table 8. Because regions do not change within cohorts, these models do not contain main effects.

The analyses reveal insightful regional differences in the effects of different labour protection measures on informal employment. In Southern European countries both employment protection (B=−.044, p<.01) and unemployment benefits (B=−.455, p<.01) significantly reduce informal work. For countries in Western Europe both unemployment benefits (B=−.229, p<.01) and social expenditure (B=−.813, p<.01) reduce the prevalence of informal work. Amongst countries in Northern Europe, only social expenditure (B=−.402, p<.1) significantly reduces informal work. In Eastern Europe, none of the labour protection measures has significant effects on informal work.

Altogether, these results suggest that labour protection measures do not have uniform effects

across European regions, and so different policy levers may work best in reducing informal work in different areas.

Table 8: Pseudo-panel linear fixed-effect model for working in the informal sector with region * labour protection interactions (all workers), model coefficients

	(22)	*(23)*	*(24)*
Eastern Europe * EPRC	0.000		
	(0.00)		
Western Europe * EPRC	-0.090		
	(-1.62)		
Northern Europe * EPRC	-0.088		
	(-1.39)		
Southern Europe * EPRC	-0.044***		
	(-3.01)		
Eastern Europe * Unemployment benefits		-0.077	
		(-1.22)	
Western Europe * Unemployment benefits		-0.229***	
		(-3.88)	
Northern Europe * Unemployment benefits		-0.046	
		(-0.61)	
Southern Europe * Unemployment benefits		-0.455***	
		(-3.41)	
Eastern Europe * Social expenditure			0.171
			(0.85)
Western Europe * Social expenditure			-0.813***
			(-4.76)
Northern Europe * Social expenditure			-0.402*
			(-1.72)
Southern Europe * Social expenditure			-0.087
			(-0.39)
N (observations)	*538*	*562*	*562*
N (cohorts)	*120*	*120*	*120*

Notes: ESS data (2004-2012). Significance levels: * = .1, ** = .05, *** = .01. Control variables as in Table 3 (except for country dummies). t-statistics in parentheses.

5. Discussion

Several mechanisms may be responsible for the observed negative relationship between labour protection legislation and informality. One channel through which this might emerge is through shifting workers' perceptions of job security (Hazans 2011a). From the worker's perspective, labour protection legislation and generous unemployment benefits incentivize the take up of formal employment (Perry et al. 2007). From the firm's perspective, labour

protection legislation incurs high financial and administrative costs. However, when enforcement of labour protection policies is reasonably strict, employers have incentives to employ workers formally as not to face penalties. These results thus support the findings in Hazans (2011a), who found that labour protection legislation was associated with lower rates of informal dependent employment in European countries in the 2004-2009 period, and contrast with those from Fialová (2010), who reports a positive association between labour protection legislation and informal work (though, as explained earlier, she considers all temporary/casual workers to be informal workers).

One reason why our results are preferable to those emerging in earlier studies is that both Hazans (2011a) and Fialova (2010) aggregate country-level longitudinal data into one observation per country and year and fit country-level fixed-effect panel regression models. This strategy is clearly inferior to our pseudo-panel models: (i) the number of country-year observations in the models and the resulting statistical power is smaller, which might give rise to Type-II estimation errors (i.e. failure to reject a false null hypothesis); (ii) their approach cannot account for compositional differences in individual-level characteristics across countries (see e.g. Duncan et al. 1998); and (iii) interpretation of this sort of models might result in 'ecological fallacies' (i.e. inferences about individual-level behaviour drawn from mean analyses) (Freedman 2002).

We show how the conclusions drawn from our preferred pseudo-panel fixed-effect models are different to those drawn from the more simple pooled cross-sectional models. Results from pseudo-panel models reveal much stronger and statistically significant associations than pooled cross-sectional models, suggesting that any similar studies using cross-sectional data and methods are likely to underestimate the negative labour protection effect on informal work. Furthermore, an inspection of the cohort-specific residuals reveals interesting patterns. The correlation between such residuals and labour protection measures is generally weak and positive, and their correlation with informal work is large and positive: cohorts exposed to high degrees of labour protection tend to have unobserved traits that are strongly associated with high levels of informal work. As a result, when the unobserved effects remain unaccounted for in the pooled models, they partly suppress the 'true' effects of labour protection on informal work.

Beyond methodological improvements, our results also provide new insights. First, our analyses cover a longer time frame than previous studies. While Hazans (2011a) and Fialová (2010) examine the 2004-2009 and 2003-2007 periods respectively, our analyses extends the observation window to the 2004-2012 period, thus engulfing the emergence and establishment of the GFC. Concerning this, we find that the inhibiting effect of employment protection on informality disappears in the GFC years, while the effects of unemployment benefits and social expenditure remain. A possible explanation is that enforcement of labour protection legislation during economic recession is lax, with court decisions leaning towards the firm, which could reduce the impact of such legislation.

Second, we compare the effects of labour protection on informal work across European regions. We find that regardless of the measure considered, when labour protection has a significant effect on the prevalence of informal work, this is always negative. Yet, our results also suggest that different labour protection measures 'work' differently across European regions. Without further inquiry, it is difficult to ascertain the reasons why this might be the case. One possibility is that this is related to the prevailing regional welfare model: the Nordic model characterised by low EPRC and high unemployment benefits, the Continental model with high EPRC and high unemployment benefits, the Mediterranean model with high EPRC and low unemployment benefits, and the Anglo-Saxon model with low EPRC and high unemployment benefits (Sapir 2006). Another possibility is that the effectiveness of different labour protection policies varies with the regional prevalence of informal work: 10% in Northern Europe, 14% in Eastern Europe, 18% in Western Europe and 29% in Southern Europe. More measures 'work' in contexts in which informal employment is more widespread. Further research into these divergences is warranted.

Third, we show that employees and self-employed workers react differently to labour protection policies. These have strong negative effect on the propensity for employees to work in formal jobs, but no apparent effect amongst the self-employed. This was to be expected, as most of the benefits of labour protection legislation – as measured here – do not extend to the self-employed (e.g. severance payments, unemployment benefits or collective dismissal regulation). Thus, this patterning of results constitutes further evidence of the

validity of our findings.[15]

Our study is nevertheless not without shortcomings. Despite the fact that our pseudo-panel models are superior to the analytical techniques used in previous studies, the reported estimates cannot be taken as 'causal effects'. Both unobserved heterogeneity and reverse causation might still operate, and so future research could attempt to move towards causal analysis by correcting these. We see two possible routes for methodological refinement in this regard, both of which are subject to data availability. Firstly, using actual panel data in which the same *individuals* are followed over time instead of pseudo-panel data in which the same *cohorts* are followed over time will attenuate any bias due to unobserved factors jointly related to labour protection legislation and informality. While pseudo-panel estimation has certain advantages over traditional panel designs (e.g. fewer issues due to attrition, measurement error or panel conditioning), mean-based pseudo-panel models do not provide information on intra-cohort effects, suffer from difficult-to-correct systematic heteroskedascity, and in some cases do not capture unobserved heterogeneity as appropriately as individual-level panel regression models (Fields and Viollaz 2013, Gardes et al. 2005). Second, reverse causality remains a possible source of endogeneity to the extent that the size of the informal sector influences the adoption of labour protection policies. Correcting for this is harder and requires the availability of an instrument or naturally-emerging experiment to be used in simultaneous-equation or instrumental-variable estimation (Wooldridge 2010).[16] Nevertheless, applications of such methods to pseudo-panel data are still in their infancy.

There are also data-driven limitations as to how *informal work* and *labour protection* are operationalized in this and other empirical studies. Survey reports of whether workers have a legal contract might suffer from *social desirability bias* (Tourangeau and Yan 2007) if

[15] Additionally, there are marked differences in the sign, magnitude and statistical significance of explanatory variables in the models for employees and self-employed workers. In fact, relatively few variables have statistically significant impacts on the propensity to work in the informal sector for the self-employed, which may be attributable to the large heterogeneity in employment circumstances amongst self-employed workers, or to selection into self-employment (Lunn and Steen 2005). Some of these differences might also be related to the measure of informal self-employment and the relatively small sample sizes for self-employed workers (particularly in pseudo-panel models).

[16] In our context, such instrument should be a factor that is moderately-to-strongly correlated with national investments in labour protection legislation, and not independently associated with informality (for example, whether a country is governed by a left- or right-wing party).

workers who do not have a contract are reluctant to admit it due to fears of being exposed or feelings of shame. This means that the observed rate of informal employment may be a lower-bound estimate of the true rate of informal employment. Relatedly, it is possible that some workers (for example the lowly educated and those who have been in their jobs for a long time) cannot recall whether they have a legal contract or are unaware of it. These issues likely introduce statistical 'noise' in the analysis, resulting in larger standard errors and more imprecise estimation of the effects of interest. Finally, despite being grounded in theory and previous empirical work, labour protection measures are imperfect, as they cannot fully capture the complexity and idiosyncrasies in legislation and enforcement across countries (Bertola et al. 2000). Better data and data collection methods would help in this regard.

Our research findings point towards promising avenues for future research. New studies should focus on empirically establishing the micro-level mechanisms that connect labour protection policies to worker and employer decisions. In this paper, we have taken a predominantly supply-side approach. From this perspective, studies should examine the reasons why different sorts of workers work in the formal and informal sectors, paying attention to the extent to which these decisions are voluntary or involuntary, and how they differ across labour protection regimes. Studies of which workers are more or less satisfied in different segments of the economy might also shed light over these issues. From a demand perspective, it is important to examine how firms react and adjust to changes in labour protection legislation in either direction. Testing this premise requires firm-level data. Finally, our analytical approach could be used to examine the relationships between labour protection and informal work in developing countries, on which systematic cross-national studies are particularly lacking (Maloney 1999, 2004).

6. Conclusion

In this paper we have analysed the associations between labour protection and informal work in European countries. We contribute to the scant existing literature by using a powerful dataset consisting of 5 cross-sections of cross-national European data augmented with external macro-level variables. Our analytical strategy – pseudo-panel fixed-effect regression models – controls for unobserved heterogeneity and compositional differences in the

characteristics of workers across countries more effectively than the specifications used in previous research. Key findings indicate that labour protection, such as stricter employment protection regulation, more generous unemployment benefits or greater levels of social expenditure, is associated with a reduction in the propensity for individuals to do informal work.

Our results have significant policy implications. In the last decade, European countries have implemented policies promoting labour market flexibility (such as reducing the generosity of unemployment benefit schemes and the regulation of employment protection) to become more competitive in the global economy (Turrini et al. 2014). This has proven effective in managing unemployment (Martin 2014). Yet, our findings suggest that these policies have the unintended consequence of increasing the size of the informal sector, which may offset (or partially offset) their alleged benefits. In this way, our analyses second Blanchard's claims that it is essential:

> *"to protect workers, not jobs […] providing unemployment insurance, generous in level, but conditional on the willingness of the unemployment to train for and accept jobs if available […] employment protection, but in the form of financial costs to firms to make them internalize the social costs of unemployment, including unemployment insurance"* (Blanchard 2006, p.45)

Of course, the desirability to invest in labour protection legislation depends also on how it affects other labour market outcomes. The tension between managing unemployment and informal work rates is not new (see Kucera and Roncolato 2008). For example, conventional wisdom dictates that labour protection may increase unemployment. This has been confirmed in empirical research (Bajada and Schneider 2009). Yet, other commentators argue that the impact of labour protection on unemployment depends on the degree to which the extra costs of labour protection are shifted onto employees through wage adjustment (Nickell 1997, p.66.). Further research taking a broader look at the pros and cons of labour protection policies is needed in this regard.

In contemporary political and media discourses the adoption of (stricter) legislation to protect workers is criticized by many on the grounds that such policies inevitably lead to increasing

informal employment. Our results stand in direct contrast to this proposition and provide important evidence that should be used to inform and contextualize these debates.

References

Acemoglu, D., Shimer R. 1999a. "Holdups and efficiency with search frictions." International Economic Review 40 (4), 827-849.

_____. 1999b. "Efficient unemployment insurance." The Journal of Political Economy 107 (5), 893-928.

Albrecht, J., Navarro, L., Vroman, S. 2009. "The effects of labour market policies in an economy with an informal sector." *The Economic Journal*, 119 (539), 1105-1129

Blanchard, O. 2006. "European Unemployment: the evolution of facts and ideas." *Economic Policy*, 21(45), 6-59.

Bajada, C., Schneider, F. 2009. "Unemployment and the Shadow Economy in the OECD." *Revue économique*, 60 (5), 1033-1068.

Bernal, R. 2009. "The informal labor market in Colombia: identification and characterization." *Desarrollo y Sociedad*, (63). 145-208

Bosch, M., Maloney, W. 2007. "Comparative analysis of labor market dynamics using markov processes: An application to informality." *IZA Discussion Paper* 3038.

Bertola G., Boeri T., Cazes, S. 2000. "Employment protection in industrialized countries: The case for new indicators." *International labour Review*, 139 (1), 57-72.

Boeri, T., Garibaldi, P., and Moen, E. 2013. "The Economics of Severance Pay." *IZA Discussion Paper*, 7455.

Cahuc, P. 2014. "Search, Flows, Job Creations and Destructions." *IZA Discussion Paper*, 8173.

Coles, M., Masters, A. 2004. "Duration-dependent unemployment insurance payments and equilibrium unemployment." *Economica*, 71 (281), 83–97.

_____. 2007. "Re-entitlement effects with duration-dependent unemployment insurance in a stochastic matching equilibrium." *Journal of Economic Dynamics and Control* 31 (9), 2879–2898.

Deaton, A. 1985. "Panel data from time series of cross-sections." *Journal of Econometrics*, 30 (1), 109-126.

Devereux, P. J. 2006. "Improved errors-in-variables estimators for grouped data." *Working Paper Series Center for Economic Research*.

Duncan, C., Jones, K., & Moon, G. 1998. "Context, composition and heterogeneity: using multilevel models in health research." *Social science and medicine*, 46(1), 97-117.

ESS, Round 6: European Social Survey Round 6 (Database 2012). Data file edition 2.1. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

_____. Round 5: European Social Survey Round 5 (Database 2010). Data file edition 3.2. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

_____. Round 4: European Social Survey Round 4 (Database 2008). Data file edition 4.3. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

_____. Round 3: European Social Survey Round 3 (Database 2006). Data file edition 3.5. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

_____. Round 2: European Social Survey Round 2 (Database 2004). Data file edition 3.4. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data.

Fredriksson, P., Holmlund, B. 2001. "Optimal unemployment insurance in search equilibrium." *Journal of Labor Economics*. 19 (2), 370–399.

Florez, L. A. 2014. "Optimal Policy with Informal Sector and Endogenous Savings." *Working paper Central Bank of Colombia* No 833,

Fields G., Viollaz, G. 2013. “Can the limitations of Panel Datasets be Overcome by Using Pseudo-Panels to Estimate Income Mobility?.” *Working paper*, Universidad Cornell-CEDLAS.

Fields, G. 1990. “Labor market modelling and the urban informal sector: Theory and evidence,” in David Turnham, Bernard Salomé and Antoine Schwarz (eds): *The informal sector revisited.* Paris, OECD, pp. 49–69.

Freedman, D. 2002. “The Ecological Fallacy.” University of California.

Fialová., K. 2010. “Labor institutions and their impact on shadow economies in Europe.” *IES Working Paper*, 29.

Gardes, F., Duncan, G. J., Gaubert, P., Gurgand, M., Starzec, C. 2005. “Panel and pseudo-panel estimation of cross-sectional and time series elasticities of food consumption: The case of us and polish data.” *Journal of Business & Economic Statistics*, 23(2), 242-253.

Hussmanns, R. (2004). Statistical definition of informal employment: Guidelines endorsed by the Seventeenth International Conference of Labour Statisticians (2003). In *7th Meeting of the Expert Group on Informal Sector Statistics (Delhi Group).*

Hazans, M. 2011(a). “What explains prevalence of informal employment in European countries: the role of labor institutions, governance, immigrants, and growth.” *IZA Discussion Paper*, 5872.

_____. M. 2011(b). “Informal workers across Europe: Evidence from 30 countries.” Background paper for In from the Shadow: Integrating Europe´s Informal Labor. Work Bank: Washington DC, *mimeo*.

Henley, A. Arabsheibani, G. R., Carneiro, F. G. 2009. “On defining and measuring the informal sector: Evidence from brazil.” *World Development* 37 (5), 992-1003.

Immervoll, H. 2012. “Reforming the benefit system to make work pay: Options and priorities in a weak labour market.” *IZA Policy Paper* 50.

ILO. 2002. “Decent work and the informal economy.” International Labour Conference, 90th Session.

Janiak, A and Wasmer, E. 2014. “Employment Protection and Capital-Labor ratios.” *IZA Discussion Paper*, 8362.

Kucera, D., & Roncolato, L. 2008. “Informal employment: Two contested policy issues.” *International Labour Review*, 147(4), 321-348.

Lonkshin, M., Ravallion, M., Sajaia, Z. 2008. “Stata command: `ppreg`. Pseudo panel regression: Fixed effect model with correction for measurement error.” Available at http://econ.worldbank.org/

Lunn, J., & Steen, T. 2005. “The heterogeneity of self-employment: The example of Asians in the United States.” *Small Business Economics*, 24(2), 143-158.

Maloney, W. F. 2004. “Informality revisited.” *World Development* 32(7), 1159-1178.

_____. 1999. “Does informality imply segmentation in urban labor markets? evidence from sectorial transitions in México.” *The World Bank Economic Review* 13 (2), 275–302.

Martin, J. 2014. “Activation and Active Labour Market Policies in OECD Countries: Stylized Facts and Evidence on their Effectiveness.” *IZA Policy Paper, 84*

MacKenzie, D.J. 2004. “Asymptotic theory for heterogeneous dynamic pseudo-panels.” *Journal of Econometric* 120 (2), 235-262.

Mortensen, D. T. 1977. “Unemployment insurance and job search decisions.” *Indus. & Labor Relations Review*, 30 (4), 505-517.

Nickell, S. 1997. “Unemployment and Labor Market Rigidities: Europe versus North America.” *Journal of Economic Perspectives*, 11 (3), 55-74.

OECD. 2013. "Employment Protection Legislation: Strictness of employment protection legislation: individual and collective dismissals", OECD Employment and Labour Market

Statistics (database). Retrieved 03 October 2014, from: http://www.oecd-ilibrary.org/employment/data/employment-protection-legislation/strictness-of-employment-protection-legislation-individual-and-collective-dismissals_data-00658-en

Pagés, C., Stampini, M. 2009. "No education, no good jobs? Evidence on the relationship between education and labor market segmentation." *Journal of Comparative Economics* 37, 387-401.

Perry, G., (Ed) 2007. "Informality: Exit and exclusion." *World Bank Publications*.

Salomäki, A. and T. Munzi. 1999. "Net replacement rates of the unemployed: Comparisons of various approaches", *EC DG ECFIN Economic Papers*, 133.

Sapir, A. 2006. "Globalization and the Reform of European Social Models." *Journal Compilation*, 44 (2), 369-390.

Schneider, F., Enste, D. 2000. "Shadow Economies: Size, Causes and Consequences." *Journal of Economic Literature*, 38 (1), 77-114.

Skedinger, P. 2010. "Employment protection legislation, evolution, effects, winners and Losers." Edward Elgar Press.

Suhrcke, M., Stuckler, D. 2012. "Will the recession be bad for our health? It depends." *Social Science & Medicine*, 74 (5), 647-653.

Tourangeau, R., & Yan, T. 2007. "Sensitive questions in surveys." *Psychological Bulletin*, 133(5), 859.

Turrini, A., Koltay, G., Pierini, F., Goffard C., Kiss., A. 2014. "A Decade of Labour Market Reforms in the EU insights from the LABREF database." *Economic Papers* 522.

Verbeek, M., Nijman, T. 1993. "Minimum MSE estimation of a regression model with fixed effects from a series of cross-sections." *Journal of Econometrics*, 59 (1), 125-136.

Wooldridge, J. M. 2010. "Econometric analysis of cross section and panel data." MIT press.

Ziegelhofer, Z. 2014. "Food prices and household welfare: A pseudo panel approach." *OCP Policy Center Working Papers*.

Zivin, K., Paczkowski, M., Galea, S. 2011. "Economic downturns and population mental health: research findings, gaps, challenges and priorities." *Psychological Medicine* (2011), 41, 1343-1348.

Appendix

Table A1: Sample means and standard deviations and t-tests

	All workers				Employees			Self-employed		
	Total	Formal	Informal	Diff. (2&3)	Formal	Informal	Diff. (5&6)	Formal	Informal	Diff. (8&9)
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*
Macro-level variables										
EPRC	2.541	2.547	2.515	0.032***	2.547	2.422	0.126***	2.546	2.562	-0.016
	(0.469)	(0.457)	(0.524)		(0.456)	(0.565)		(0.465)	(0.496)	
Unemployment benefit	0.640	0.647	0.612	0.035***	0.647	0.593	0.054***	0.644	0.621	0.023***
	(0.161)	(0.155)	(0.184)		(0.155)	(0.204)		(0.164)	(0.172)	
Social expenditure	0.232	0.234	0.228	0.006***	0.234	0.221	0.013***	0.236	0.231	0.006***
	(0.041)	(0.042)	(0.037)		(0.042)	(0.036)		(0.041)	(0.038)	
Per capita GDP growth	0.038	0.039	0.035	0.004***	0.039	0.034	0.005***	0.036	0.036	0.001
	(0.035)	(0.035)	(0.036)		(0.035)	(0.038)		(0.036)	(0.036)	
Micro-level variables										
Self-employed	0.139	0.033	0.665	-0.632***						
Female	0.478	0.496	0.395	0.102***	0.502	0.513	-0.010	0.313	0.335	-0.022**
Age	42.29	41.81	44.64	-2.828***	41.63	40.98	0.645***	47.17	46.47	0.694***
	(12.15)	(11.88)	(13.12)		(11.84)	(14.52)		(11.93)	(11.94)	
Partnered	0.678	0.677	0.683	-0.006	0.675	0.574	0.101***	0.753	0.738	0.015
With children	0.498	0.499	0.498	0.001	0.498	0.429	0.070***	0.504	0.533	-0.028**
Disabled	0.004	0.004	0.005	-0.000	0.004	0.006	-0.001	0.005	0.004	0.001
Lives in a rural area	0.073	0.062	0.132	-0.070***	0.061	0.119	-0.057***	0.068	0.138	-0.070***
Student	0.036	0.036	0.037	-0.001	0.037	0.081	-0.045***	0.024	0.015	0.010***
Retired	0.016	0.014	0.032	-0.018***	0.013	0.034	-0.021***	0.040	0.031	0.009**
Education										
Primary	0.059	0.046	0.126	-0.079***	0.047	0.140	-0.093***	0.030	0.119	-0.089***
Secondary	0.618	0.614	0.642	-0.028***	0.623	0.635	-0.012*	0.361	0.646	-0.284***
Tertiary	0.319	0.338	0.230	0.108***	0.329	0.223	0.106***	0.604	0.233	0.371***
Occupation skill level										
Low	0.452	0.425	0.585	-0.160***	0.437	0.637	-0.200***	0.081	0.559	-0.478***

Medium	0.266	0.286	0.172	0.114***	0.294	0.190	0.104***	0.046	0.163	-0.117***
High	0.280	0.289	0.243	0.046***	0.269	0.172	0.096***	0.873	0.278	0.595***
Local minority	0.018	0.019	0.018	0.001	0.019	0.017	0.002	0.017	0.019	-0.002
Immigrant	0.050	0.050	0.055	-0.005***	0.050	0.088	-0.038***	0.043	0.038	0.005
Employment sector										
Agriculture	0.045	0.026	0.140	-0.114***	0.026	0.069	-0.044***	0.025	0.175	-0.150***
Manufacturing	0.159	0.173	0.095	0.078***	0.176	0.122	0.054***	0.099	0.082	0.017***
Construction	0.297	0.265	0.452	-0.186***	0.263	0.443	-0.179***	0.325	0.456	-0.131***
Transport and trade	0.188	0.190	0.181	0.009**	0.185	0.150	0.035***	0.325	0.197	0.128***
Public services	0.309	0.346	0.132	0.214***	0.350	0.216	0.134***	0.225	0.090	0.136***
Observations	*93,948*	*77,982*	*15,966*		*75,442*	*5,355*		*2,540*	*10,611*	

Notes: ESS data (2004-2012). Standard deviations in parentheses. Significance levels for t-tests of mean differences: * = .1, ** = .05, *** = .01.