Risk-sharing and the Decentralization of Social Insurance

Pablo Beramendi
Department of Political Science
Duke University

Prepared for presentation at the workshop on Federalism and Decentralized Governance, Stanford University, July 21-22 2008

-Preliminary Draft. Comments welcome-

In developing this paper I have benefited from comments from and/or conversations with Carles Boix, Alberto Diaz-Cayeros, Daniel Kselman, Suzanne Mettler, Jonathan Rodden, David Rueda, David Soskice, and Erik Wibbels.
Abstract

The emergence of unemployment insurance in North America poses a puzzle for comparative political economists. Though departing from very similar institutional and economic circumstances, the Depression yielded very different responses in the two North American federations. While Canada developed a centralized system of unemployment insurance in 1941, the Social Security Act (1935) included provisions leaving the ultimate design and implementation of unemployment insurance in the hands of the states. Why? To address this question I put these two experiences in the context of a more general argument about endogenous fiscal institutions. I argue that, given a set of decision-making rules and procedures, economic geography, most notably interregional differences in income and labor market risks, and its interplay with the level of crossregional economic externalities drive the type of coalitions emerging in support of different levels of centralization of social insurance. I show how the balance between these different elements varied substantially between Canada and the United States, thereby explaining why they ended up adopting different responses to the Great Depression.
The emergence of unemployment insurance in North America poses a puzzle for comparative political economists. Unemployment insurance programs emerged as a response to the economic shock that shuttered the social, economic, and political world on both ends of the Atlantic. Interestingly enough, the Depression yielded very different responses in the two North American federations. While Canada developed a centralized system of unemployment insurance in 1941, the Social Security Act (1935) included provisions leaving the ultimate design and implementation of unemployment insurance in the hands of the states. Though eventually transformed into a hybrid regime (Lieberman 1998; Mettler 2002), the system remained one in which states can differ in terms of pre-requisites, generosity and duration of unemployment benefits. Throughout most of the 20th century this has been one of the most prominent small differences that matter to explain why distributive outcomes differ between the two North-American federations (Card and Freeman 1993; Banting 1997: 267-309; Myles 1998:341-364). This diversity of responses is particularly puzzling because both unions confronted the Great Depression from very similar positions.

During the first three decades of the XXth century Canadian and American social welfare systems shared many features. Among these stand out the endorsement of the British Poor Laws system and its underlying philosophy and their political and organizational fragmentation. In both countries, with some remarkable exceptions, the worse off were left to be taken care of by way of charity, either organized by the local administration or directly dispensed by private philanthropic organizations. Put it simply, the ruling approach to the indigents was that they did not constitute a public policy issue. Those who were laid off were expect to fend for themselves and search for whatever jobs were available. Only those lacking any possible opportunity (because of illness or age) were considered the target of the relief provided normally by locally organized voluntary associations and only exceptionally (civil war pensions) by the federal government (Skocpol 1992). In general during the first fifteen years of the century public authorities were not directly concerned with social issues. State/provincial and federal governments refrained from direct intervention and only in exceptional circumstances did they launch one-off programs of emergency assistance.

Political fragmentation was an attendant feature to the absence of involvement of public authorities in poor relief and redistribution. In the United States the use of local poor relief was heavily linked to the organizational features of the patronage democracy (Orloff 1988). Regarding Canada, as Pal puts it, "the British North America Act had little to say about those activities that now comprise the welfare state [...]” In addition, both in Canada and the United States at the time, political representation follows a decentralized form of federalism (Chhibber and Kollman 2004; Finegold and Skocpol 1995). Thus, regional elites have the institutional capacity to prevent the federal government from launching any program potentially disruptive of their specific political economy. Regarding unemployment insurance, Canadian provinces had veto over any amendment altering the distribution of powers across territories of the British North America Act. In the case of the USA, state rights were also at play, protected by the Supreme Court. In addition, Southern representatives enjoyed a privileged position in both the House and
the Senate. The lack of political rights of the black majority of the labor force generated an effective single party dominance of the Democratic Party (Key 1949; Kousser 1974; Alston and Ferrie 1999). This, in turn, gained them seniority in both houses. And with seniority came the capacity to determine the agenda in the relevant Committees, which implied a huge shaping power of any legislation under consideration. If the status quo and the Depression affecting the economy were similar, if political representation was similarly biased towards states and provinces, why did the paths of these two rather similar systems diverge in between the late 1930s and early 1940s?

To address this question I put these two experiences in the context of a more general argument about endogenous fiscal institutions. I argue that, given a set of decision-making rules and procedures, the relative importance of different dimensions of economic geography, most notably interregional differences in income and labor market risks, and their interplay with the level of interregional economic externalities drive the type of coalitions emerging in support of different levels of centralization of social insurance. I show how the balance between these different elements varied substantially between Canada and the United States, thereby explaining why they ended up adopting different responses to the Great Depression despite the large similarities in their pre-existing institutional conditions.

I proceed as follows. First, I present a theoretical analysis of the links between economic geography and institutional preferences. From this analytical perspective, section two focuses on the nature of the Great Depression and its implications for the patterns of economic geography. Sections three to five trace the political process leading to two different institutional solutions to the demand of introducing a national, comprehensive, unemployment insurance program. Finally, I conclude with a discussion of the main theoretical implications emerging from the analysis.

I. The Politico-economic Model.

Confronted with a change in the economic environment, political actors must choose whether to adjust existing fiscal structures. The options available for citizens and politicians are FC (both interpersonal and interregional redistribution are provided by the national executive), FI (independence/ decentralization, i.e., interpersonal redistribution is provided by the regions with no interregional transfers), and HR (systems dominated by large levels of horizontal redistribution between regions and decentralized interpersonal taxes and transfers). Thus, a change in the direction (FI→HR→FC) is a move from fragmented solidarity to fiscal integration, whereas a change in the direction (FC→HR→FI) is a move from a centralized to a decentralized system of redistribution.

In any given union with several levels of government, this process of adopting this decision is driven by the interplay between local and national elites. To simplify matters, I assume that the party systems at the local and national levels are similar. That is, the parties (P) competing across regions (indexed by r=1…r) and the national level (N) are the same. Local elites value office holding in their regions but also derive some benefit from their party winning the national election. Likewise, national elites gain
utility from holding the national executive as well as from having their party win regional elections. Accordingly, their respective utility functions are defined as follows.

The utility of the any given party, P, leader in a district (r) is given by:

\[ U_r^P = p_r^P \lambda_r^P + \Phi(q_n^P \phi_r^P) \]  

(1)

where \( p_r^P \) defines the probability of party P conquering the regional executive, and \( \lambda_r^P \) represent the rewards obtained by regional party elites from holding office at the national level. The second half of the expression captures the benefits that regional party leaders obtain from having their national elites win the national elections. Thus, \( q_n^P \) defines the probability of party P conquering the national executive and \( \phi_r^P \) captures the utility gains derived from having the same party control the national executive.

In turn, the utility the of P’s national leadership is captured by:

\[ U_n^P = q_n^P \lambda_n^P + (1 - \Phi) \sum (p_r^P \chi_r^P) \]  

(2)

where \( \lambda_n^P \) represent the rewards obtained by national party elites from holding office at the national level and \( \chi_r^P \) captures the benefits that the national elite of party P derives from winning elections in region r.

This framework points to two constraints on the behavior of political elites that are central to the development of a theory of fragmented solidarity. The first constraint is institutional. \( \Phi \) is a parameter that captures the importance that party elites give to the result of elections held at a level of government other than their own. This parameter ranges between 0 and 1. A high level of \( \Phi \) indicates that national elections are of great importance in local leaders calculations, whereas low values of \( \Phi \) (and consequently, high values of \( 1 - \Phi \)) indicates that the results of regional elections are an important concern for national elites. The second constraint is electoral as, under democracy, politicians are bounded to a large extent by their constituencies’ preferences. Critically, the probabilities of national and local elites to gain office reflect the need by both the national and the regional elites to minimize the distance between their proposals and citizen’s optimal institutional preference. Let \( X_r^* \) and \( X_n^* \) represent the optimal institutional preference of the pivotal group of voters at, respectively, the regional and national level of government. In turn, \( X_r^P \) and \( X_n^P \) define the proposals made by the regional and national elites of party P. By assumption, the distance between voter’s ideal points and elites proposal are constrained to be between 0 and 1. Thus, the probability that the national elite will win the national election \( (q_n^P) \) is given by

\[ 1 - (X_r^* - X_n^P) \leq 1 - \Omega_n \] under fiscal decentralization and by \( 1 - (X_n^* - X_r^P) \leq 1 - \Omega_n \) otherwise. Similarly, the probability that the regional elite wins the regional election \( (p_r^P) \) is given by \( (1 - (X_r^* - X_r^P)^2) = 1 - q_r \) under decentralization and by

\[ (1 - (X_r^* - X_n^P)^2) = 1 - q_n \] under centralization.
Within this analytical framework, Figure 1 displays the two possible patterns of strategic interactions between the two levels of government, as determined by the status quo. Under SQ1 the distributive conflicts associated with the geography of inequality emerge in a context of fragmented solidarity. Under these circumstances, the national elites evaluate whether to pursue centralization. If they opt not to, the game ends where it began. National elites operate under uncertainty. When they evaluate the pursuit of centralization, they are not sure what the response of the regional elites will be. Should national elites launch a proposal to move from a system of fragmented solidarity to a system of fiscal integration, regional elites will choose at $t+1$ between accepting (with probability $\alpha$) or rejecting such move (with probability $1-\alpha$). After the decision by the regional elites, the game ends either with an institutional change (FS $\rightarrow$ FI) or with the preservation of the status-quo in the aftermath of a conflict between national and regional elites. As a result of such conflict, the electoral chances of the national government ($q^*_n$) are reduced by a factor of $\frac{1}{\varepsilon}$ with $\varepsilon > 0$. The sequencing of the game has a similar structure under SQ2, that is to say when distributive tensions become politically relevant in a context of fiscal centralization. The mover at $t$ is now the regional elite, that must decide whether to pursue fiscal decentralization (FI $\rightarrow$ FS). As before, it does so with some uncertainty about the response by the national elite at time $t+1$ (as captured by probabilities $\beta$ and $1-\beta$). A potential conflict between regional and national elites also reduces the electoral chances of the first mover ($p^*_r$) by a factor of $\frac{1}{\varepsilon}$.
Solving the model requires understanding the incentives of first movers to challenge the status quo given the expected response of other party level. In equilibrium, the indifference conditions for the first movers yield the following expressions:

\[
\frac{(1 - \Omega_r)}{\Omega_n} = \alpha \frac{\left[ \frac{1}{\varepsilon} \Omega_r + \frac{(1 - \Phi) \chi_r (\varphi_r - \varphi_n)}{\lambda_r^p} \right]}{\Omega_n (1 - \frac{1}{\varepsilon})} \quad \text{for SQ1} \tag{3}
\]

and

\[
\frac{(1 - \varphi_n)}{\varphi_r} = \beta \frac{\left[ \frac{1}{\varepsilon} \varphi_n + \frac{\Phi \phi_r (\Omega_n - \Omega_r)}{\lambda_r^p} \right]}{\varphi_r (1 - \frac{1}{\varepsilon})} \quad \text{for SQ2} \tag{4}
\]

The following insights emerge from expressions (3) and (4).

First, the probability of agreeing to an institutional change increases in the electoral price of intra-party conflict. As a result, central governments are more likely to yield to the institutional demands of particularly relevant regions within the country. Indeed, \(\varepsilon\) under SQ2 can be thought of as an indicator of the bargaining leverage of certain regions within the coalition sustaining the party nationally. In turn, under SQ1 high values of \(\varepsilon\) reflect higher levels of dependency by regional elites on the resources facilitated by the national leadership of the party.

Second, as the value attached by regional leaders to the incumbency of the national executive increases, the probability that they accept a change towards higher levels of fiscal integration also increases (\(\frac{\partial \alpha}{\partial \Phi} > 0\)). In contrast, as the importance of national political contests rises, the probability that the national party elite agrees to decentralize redistribution declines (\(\frac{\partial \beta}{\partial \Phi} < 0\)). Building on the seminal contributions by Crémer and Palfrey (1999, 2000), the parameters capturing the electoral cost of cross-jurisdictional conflict (\(\varepsilon\)) and the balance between the interests of national and local elites (\(\Phi\)) define the political playing field in the event of increasing distributive tensions among territories and income groups. A key distinction here is whether the way coalitions are formed is centrifugal, that is tailored towards the representation of territorialized interests, or centripetal, that is set up to increase the incentives of political actors to mobilize the interests of social groups that cut across subnational territorial boundaries. Centrifugal representation corresponds to very low values of both \(\varepsilon\) and \(\Phi\). In contrast, centripetal representation implies the preeminence of political competition at the national level (high values of \(\Phi\)) and very high electoral costs for conflictive regional elites (\(\varepsilon\)).

Third, distributive conflicts associated with economic geography work to drive elite’s preferences apart through their impact on the electoral constraints they face. As
these tensions exacerbate, the probability that either level of government responds favorably to the possibility of institutional change becomes smaller and smaller. At the extreme, when this probability gets sufficiently small, first movers find no incentive to challenge the status-quo.\(^1\) The role of economic geography in driving the electoral constraints on political elites is particularly central to understand the different institutional choices between Canada and the United States because, as mentioned above, both countries have similarly centrifugal systems of representation in the 1930s (i.e. low values of \( \Phi \) and \( \varepsilon \)). Thus, in accordance with the historical circumstances of the cases under analysis, I proceed on the assumption that regions are veto players that can accept or reject any proposal to centralized unemployment insurance.

To analyze the origin of institutional preferences, consider a union with just two regions (A, B). A is poorer than B, i.e. it has a lower aggregate income per capita \((w^A < w < w^B)\). Citizens in both regions vary in their level of pre-tax income \((w_i)\) and are driven both by redistributive and insurance concerns (Varian 1980; Atkinson 1995; Moene and Wallerstein 2001). This being the case, risks, aversion to them, and their territorial distribution become a primary mechanism for other aspects of economic geography to condition the selection of fiscal institutions. Risks are assumed to behave according to Arrow-Pratt constant relative risk aversion function, by which \(\left( \frac{c_i}{1-\delta} \right)^{1-\delta}\),

where \(c_i\) represents individual consumption. At any given time individuals might be employed with probability \(\alpha\) or unemployed with probability \((1-\alpha)\). Individuals maximize consumption across both states. Individual consumption is defined by \(c_i = (1-t)w_i\) in the good state of the world and by \(b\) in the bad state of the world, where \(b\) represents the benefits individuals receive while being unemployed. In addition, citizens are affected by an interregional transfer that, when in place, is a function of the regional average income vis-à-vis the union. Thus, citizens face a decision about two policy instruments, namely the level of interpersonal (vertical) redistribution \((t)\), and the level of horizontal transfers of resources among the members of the union, that is to say the level of horizontal redistribution \((T)\) according to the following utility function:

\[
V_i = \alpha \frac{c_i^{1-\delta}}{1-\delta} + (1-\alpha) \frac{b^{1-\delta}}{1-\delta} - T(w - w^u)
\]

s.t. \(\alpha w t = b(1-\alpha) + T(w - w^u)\)

Solving the citizens’ optimization problem under these circumstances yields:

\(^1\) Ceteris paribus, as the institutional preferences associated with income converge, that is as \((q_r - q_n)\) tends to zero, the probability \(\alpha\) that the regional elite accepts a centralization of the fiscal system increases (and vice-versa). Likewise, as the gap between the national and the regional distributions of income increase \((\Omega_n - \Omega_r)\), so do the electoral costs of decentralizing redistribution for the national elite.
\[
\frac{1 - t_i - T}{t_i} = \left( \frac{w_i}{w} \right)^{\frac{1}{\delta - 1}} \frac{\alpha}{1 - \alpha}
\]

(6)

A number of interest implications follow from this expression:

(1) The demand for redistribution increases with the scope of realized risks. That is to say, the preferred level of redistribution increases in the probability of unemployment and, conversely, declines with the probability of employment. This is captured by the fact that \( \frac{\partial t}{\partial \alpha} < 0 \).

(2) The demand for redistribution and insurance increases with risks aversion. To see this, note that in the solution to the optimization problem, it is the case that \( \frac{\partial t}{\partial \delta} > 0 \).

(3) For high levels of risk aversion, it is the case that the demand for redistribution increases with income. This result also follows from (6) in that

as \( \delta \to \infty \), then \( \frac{\partial t}{\partial w_i} > 0 \)

(4) Finally, expression (6) also illuminates citizens’ preferences for interregional redistribution by showing how the resources obtained (lost) through horizontal redistribution affect directly any entity’s ability to engage in interpersonal redistribution (t). As a result,

\[
T = \begin{cases} 
0 & \text{if } w \geq w^u \\
0 < T < 1 & \text{if } w < w^u 
\end{cases}
\]

That is to say, citizens in regions poorer than the union (\( w < w^u \)) benefit from the existence of interregional transfers of resources, whereas those in regions wealthier than the average (\( w > w^u \)) would suffer from it.

These results shed light on the linkages between the geography of inequality and labor market risks and the institutional preferences of voters. Figure 2 displays the analytical results of the model when risks associated with regional economic specialization affect either the poorer or the wealthier member of the union.
Consider first the role of income inequality between regions. This is best captured by looking at a union where workers in both regions have similar levels of risk aversion. Thus, assume a new country emerges out of merging region B in panel 2.a with region A in panel 2.b. Interestingly, none of the four groups in this hypothetical union share a common set of preferences over the two policy instruments (t, T). This is a reflection of the interplay between the redistributive conflict among individuals within regions and the redistributive conflict between regions within income groups. Poorer citizens in the poor region want to maximize redistribution through a fully centralized system where both interpersonal and interterritorial redistribution are high. In contrast, wealthier citizens in the rich region want to preserve fiscal independence, thus minimizing redistribution on both dimensions. The other two groups show conflicting motives. The rich among the poor, i.e. the elites of poorer regions, want to maximize interregional redistribution to increase their resources and minimize interpersonal redistribution. In turn, the poor among the rich want to increase interpersonal redistribution without sharing their tax base with poorer citizens of any other region. Given the ranking of preferences of the different actors involved, and the fact that provinces have, as inequality between regions increases, the political feasibility of centralized public insurance policies declines.

The distribution of preferences changes in the preferences of cross-regional differences in the level of economic specialization. According to the model’s results, the elites of highly specialized regions have incentives to preserve their own system of interpersonal redistribution, even though it brings about some level of taxation on their income. In turn, labor in asset specific regions is less likely to move, which makes it more risk averse and likely to demand more insurance (Boix 2003). Accordingly, differences about the value of interpersonal redistribution (t) among income groups decline as a result of economic specialization. Therefore, the formation of a political coalition in support of regional interests becomes more likely. Interestingly, this dynamics is orthogonal to the level of wealth of different regions: with sufficiently high levels of
regional risk specificity, the citizens of a region poorer that the union may be willing to forego the income benefits of centralizing interpersonal redistribution to protect its ability to design an insurance system better tailored to their specific risk profiles. By reducing the levels of cross-regional risk sharing, differentials associated with the diversity of economic activities across regions work to facilitate the formation of intra-regional cross-class coalitions, and, conversely, to undermine the likelihood of inter-regional class based coalitions. To the extent that political representation is centrifugal, as it in the case in both Canada and the United States during the 1930s, a diverse geography of risks will work to preserve a decentralized system of interpersonal redistribution (t), thereby transforming political contentions towards a conflict over the scope horizontal redistribution (T).

Despite its powerful imprint, an uneven economic geography is not the only factor shaping the levels of risk-sharing between regions. For one, external economic shocks are a source of risk that fosters the political demand for a common pool of resources, that is to say in favor of relatively more integrated fiscal structures (Alesina and Perotti 1998; Cremer and Palfrey 2000). Neither the poor nor the wealthy region know ex ante whether or how they will be affected by a shock. What they both know, though, is that if a negative shock actually hits them, they would be worse off in the absence of a common insurance scheme. Consider the economic effects of a negative shock on regions A or B. Any region affected by a negative shock faces simultaneously a loss of tax base (w) and an increase in the demand for redistribution (via the ratio between the employed and the unemployed populations, $\frac{\alpha}{1-\alpha}$). To the extent that there is a cross-regional fiscal structure at work, the region negatively affected by the shock can transfer some of the cost to the common pool. Otherwise, it must fend for itself. Therefore, the possibility of an external shock creates a risk that cuts across regional boundaries, thereby compensating the effect of specific risks associated with the uneven geography of labor markets. As a result, they create incentives for the formation of coalitions in support of a more integrated fiscal system. The question of which specific form this fiscal system should take follows naturally, that is whether the adopted fiscal structure will be fully centralized (FC) or will be one dominated by interregional, horizontal redistribution (HR). In addressing this question, the key issue is whether the negative socio-economic effects associated with the common shock span across regions or remain geographically concentrated.

In the event of a negative shock affecting their local economy, it is only natural for laid-off workers to seek alternatives in those areas of the country less affected by the crisis. At the same time, in the absence of a centralized welfare system, the unemployed and their dependents will also seek the welfare offered in other localities (Peterson 1990; 1995). The geographical mobility of the unemployed and their dependents has several effects. On the one hand, it reduces the distance across regions in terms of the tax base (w) and the size of the dependent population ($1-\alpha$). The region expelling unemployed, poor people increases its employment rate ($\alpha$) and average output (w), whereas the recipient region sees both magnitudes drop. As a result, both regions come closer in terms of the nature of the distributive conflict among their citizens. As the poor travel across
regional boundaries, the ability of the poor in wealthier regions to protect its tax base by reducing horizontal redistribution (T) becomes fruitless, and the incentives to support a more centralized fiscal structure (FC) increase. By fostering the levels of regional risk-sharing, externalities work to reduce the terms \((q_1 - q_n)\) in SQ1 and \(\Omega_n - \Omega_r\) in SQ2, thereby undermining the incentives of regional elites to establish their preferred fiscal structures even under conditions of centrifugal representation that grant these elites effective veto power. Indeed, for sufficiently large levels of regional-risk sharing, the very incentives of regional leaders to even challenge a centralized fiscal system decline decisively. This interdependency across regions is captured in the left panel of figure 3: interregional mobility of dependents from the economically depressed to the economically prosperous areas implies by definition an horizontal transfer of resources between the regions \((T>0)\) of the union. As the policy preferences of the different groups grow closer, the political pursuit of a fiscally centralized system becomes more feasible.\(^2\)

\[\text{Figure 3: Policy Preferences in the presence of economic externalities across regions.}\]

The comparison between the left and right panels in figure 3 also helps understand the conditions under which a coalition in favor of a fully centralized fiscal system is more likely to emerge as a result of economic externalities associated with the mobility of dependents. A requirement for this to occur is that the levels of economic specialization be similar across the different regions of the union. Only then the different groups involved find a common ground about interpersonal redistribution (t) in a context in which large levels of interregional redistribution are effectively at work because of mobility itself.

In contrast, the distribution of preferences is more heterogeneous when large levels of mobility among dependents combines with differential patterns of regional economic specialization (panel b). While citizens from the poor, non specialized region would have

\[\text{For other analyses where increasing labor mobility facilitates the adoption of common social policies, see Wildasin (1991, 1995), Bolton and Roland (1996) and Perotti (2001).}\]
a lot to gain from a fully centralized system, citizens in the wealthy, economically specialized region face a complex set of motives that work in opposite directions. On the one hand, they need some form of common pool of resources against the prospect of a external negative shock. Yet, at the same time, they want to preserve policy autonomy to design interpersonal redistribution in ways that disrupt the least the workings of their regional economy. Due to high levels of regional economic specialization ($\delta$), the possibility of a large contingent of dependent immigrants affects not only the regional tax base, but also the existing fit between decentralized redistributive efforts and the organization of local labor markets. As a result, those who have more to fear from the larger levels of vertical redistribution spurred by large levels of dependents’ mobility have incentives to act strategically.

That is to say, the elites of the wealthier region, in anticipation of undesired population inflows, have strong incentives to help fund decentralized redistribution in other regions through larger horizontal transfers (T). Citizens of the wealthier region (both rich and poor) coalesce with the rich citizens of the poor region around an exchange in which horizontal transfers (T) are used to limit the extent of interpersonal redistribution (t), even if at the expense of the poorest members of society. That way, they limit the scope of migrations that would undermine the viability of more decentralized inter-personal redistribution. This will be the optimal strategy until an additional marginal increase in horizontal redistribution T equates the net loss due to changes in interpersonal redistribution (t) motivated by larger levels of cross-regional mobility of dependents. In conclusion, in the presence of conflicting motives, such as large levels of dependents mobility on the one hand, and large levels of regional economic specialization on the other, support for a combination of lower levels of vertical redistribution (t) and larger levels of horizontal redistribution (T) are expected to grow stronger. The resulting fiscal structure would be one dominated by interregional redistribution (HR). The underlying logic here is one of prospective self-insurance, one that is also at work in the international arena: the relatively privileged rather pay some amount to keep the poor away than facing the risk of having to incorporate an undesired contingent of dependents into their economies.

To summarize, given centrifugal representation and an uneven geography of inequality and insurance,

1. The centralization of public insurance systems (FC) occurs when large interregional economic externalities are at work.

2. Fiscal structures dominated by interregional redistribution (HR) are expected to emerge when, in a context of centrifugal representation, large geographical disparities in terms of income and labor market profiles coexist with large economic externalities. Under such conditions, the leaders of highly specialized regions transfer resources strategically to protect their own regional economies from undesired population movements.
I turn now to examine whether these predictions shed light on the origins of the institutional divide in unemployment insurance among the North American federations.

II.- The Great Depression and the pressures for institutional reform.

Canada and the USA were hit similarly by the sequence of economic downturns that reached its peak during the late 1920s and early 1930s. Briefly mentioned, these were four: an ongoing process of deruralization that created a massive surplus in the Canadian and American labor forces, the European monetary crisis (1930-31), an insufficient and late reaction in terms of macroeconomic policy, and finally and most visibly, the worldwide collapse of financial and stock markets. Together these four factors provoked a long lasting reduction in both wealth and consumption, themselves affecting the expectations about recovery and thus making the Depression longer (Temin 1976: 62-96:138-179). The social consequences the Depression were dramatic on both sides of the frontier. Figures 4 and 5 display both the magnitude as well as the timing of the Depression through two indicators: the percentage of the total civilian labor force that is "not working and seeking for a job"\(^3\), and the GNP per capita.\(^4\)

---

\(^3\) Sources: Canada. Historical Statistics. USA: the unemployment rate is defined as one minus the proportion of the civilian labor force that is actually employed (Historical Statistics of the United States. Colonial Times to 1957, U.S. Department of Commerce, Bureau of the Census, 1961, pp 68-69).

\(^4\) These are the Gross National Product per capita at market prices, the Gross National Product at 1929 prices and the Disposable Income per capita at market prices. Source: see previous fn). Canada: Historical Statistics and represent Gross National Product at Market Prices.
Issues of comparability aside, the trend shown by these indicators is similar in both countries. The peak in the unemployment rate is 1931, with 19.32% of the civilian labor force in Canada and around 24% in the USA. Likewise, the GNP per capita also hit the floor during the period 1930-1931 in both nations. Talking about the USA, Achembaum (1986: 16-17) offers the following picture of the implications of the Depression:

"[...] this time the extent, intensity, and duration of the upheaval were unprecedented. Between October 1929 and June 1932, the common stock price index dropped from 260 to 90. The nation’s real GNP, which had risen 22 per cent between 1923 and 1929, fell 30.4 percent over the next four years. Nearly 5000 banks, with deposits exceeding $3.2 billion, became insolvent; 90000 businesses failed. The gross income realized by farmers was cut nearly in half.[...] More than one thousand local governments defaulted on their bonds[...] Breadlines formed[...] Bankrupt firms could not honor their pensions obligations to superannuated workers.[...] Misery was a threat to everyone"

The picture looked no better in Canada (Struthers 1983: 44-104). The Royal Commission on Dominion-Provincial Relations put it bluntly when stating that "the livelihood for hundreds of thousands of citizens seemed to be entirely dependent upon public charity" (1940: 162). Along with its widespread scope, the Depression also had a geographical dimension. Figures 6 taps into this dimension by mapping the distribution of states and provinces along three variables: the level of regional income per capita at the beginning of the crisis (1929), the drop in personal income per capita between then and the Depression’s peak (1934-35), and the level of financial capacity of subnational governments, as captured by the state/provincial revenues per capita in 1930.

Amidst a context of economic decay (after all, the “lucky” provinces/states saw drops of 30-40% in their levels of income per capita), it seems undeniable that the Depression hit some regions much harder than others. Those states/provinces whose economies were of an agrarian basis were especially damaged by the fall in consumption

---

5 D.Guest (1997:83) offers estimates that put Canada closer to the USA: "At the bottom of the Depression in 1933 nearly one quarter of the labor force was out of work and seeking jobs and an estimated 15 per cent of the population was in receipt of relief".
and the attendant drop in agricultural prices.\textsuperscript{6} In Canada these were the western, prairie provinces, Alberta and Saskatchewan most prominently.

**Figure 6: The Geography of the Depression in the United States and Canada**

As Figure 6 displays these two provinces not only were among the poorest in terms of income per capita in 1929 but also took the worst hit, together with Prince Edward Island, by losing between 60 and 70 percent of their income five years into the Depression. In turn, the upper left quadrant (i.e. those with income levels below average and income losses above average) in the United States is populated by the Midwest and Southern States, where a lot of planters were unable to keep their labor force. In addition to the fact that these states had very few resources at their disposal (their revenues per capita were scant at best), charitable institutions had very little, if any, presence in rural and agricultural dependent areas (Patterson 1986). As a result, during these years chronic, long-term poverty is to be seen as particularly acute in the South (Alston and Ferrie 1999: 49-50). Generally, it was the case in both countries that those with the worst problems had the least resources with which to respond. The situation was particularly extreme

\textsuperscript{6} Guest (1997:83-135) reports, for instance, a fall in the price of wheat from $1.60 in 1929 to 38 cents in 1931
given the very low fiscal capacity for states and provinces: while the average provincial income per capita in Canada was around $240, the average provincial revenue per capita was $18. Similarly, while the average state income per capita was above $600, the average state revenue per capita was just above $20. The right hand side of Figure 6 also suggest that this scarcity was quasi-uniform throughout Canada, with the relative exception of British Columbia, and slightly more heterogeneous in the United States. Faced with such a worsening of social conditions the existing welfare institutions were politically and financially powerless.

These patterns fostered distributive tensions within and between territories, and the demand for adjustment of existing fiscal and social security institutions. Bankrupted provinces and states plead Washington and Ottawa for help. Others, for different reasons to be explored below, were wary of too much federal intervention. I turn now to analyze how these tensions resolved in each case.

III.- The Canadian Response: the Centralization of Unemployment Insurance.

The Unemployment Insurance Act (1940) is the result of a political process that reverted the existing territorial design of welfare provision since the origins of the country. Its approval took a major constitutional amendment that spanned between 1936, a year in which the provisions regarding unemployment of the Bennett’s New Deal (1935) were declared unconstitutional by the Supreme Court of Canada and the Privy Council of Great Britain, and 1940, a year in which the Federal Government and the provinces agreed to reform the British-North American Act so that the former could, among other things, take full control of the emerging national program of Unemployment Insurance. However, the struggles about unemployment started in Canada much earlier than 1935.

To a large extent the history of the emergence of the Unemployment Insurance Act (1940) is the history of a failure, namely the ultimate failure to preserve the welfare provisions established under the British North America Act (1867), in particular its resort to local and provincial relief and charities as the optimal layer to deal with a problem, unemployment, that was viewed as an individual issue. The unemployed were so because "some fault of their own" and hence society should apply the criteria of "less elegibility", which implies caring just for those obviously unable to fend for themselves (e.g. war injured). Such a view was shared by Liberals and Conservatives during the 20s and 30s.

Yet, as economic conditions worsened and governments muddled through recurrent downturns, both the demands by provinces and municipalities and the stringency of public finances grew stronger. Unemployment became central to the political agenda, conditioning both the electoral outcomes of the competition between the Liberals (1921-1930; 1935-1941) and the Conservatives (1930-1935) and the dynamics of the Dominion-Provincial relations. Much of the contention was about passing the buck. As the Provinces and the Municipalities were overloaded, Ottawa tried to keep its financial and administrative involvement to a minimum and hence insisted on
the fact that, constitutionally, taking care of the unemployed was their task. Ottawa remained committed to this view and succeeded until the second half of the 1930s. By 1940 the roles were exchanged: unemployment insurance was designed as a fully federal program, national in scope. To understand the process driving this radical change it is useful to distinguish three periods: the 1920s, especially the second half, with the first Liberal Government of Mackenzie King; the first half of the 1930s (1930-35), with Bennett’s Conservative Government and, finally, the period 1935-1940, when the Liberals of Mackenzie King return to power.

During the early 1920s the prospects for a national unemployment insurance policy looked promising. In 1919 the Royal Commission on Industrial Relations recommended a national insurance schema to deal with the temporarily unemployed (Mcintosh and Boychack 2001:81-82). To cope with the demobilization after World War I an Employment Service had been established, collecting for the first time systematic information about the labor force and re-allocating workers throughout the country. Moreover, the officials of the Department of Labor recommended the adoption of a system of unemployment insurance. The context changed rapidly half way through the decade: a slight economic recovery put the issue aside in the political agenda. More importantly, the Liberals were in power in coalition with the Progressive Party, an organization created to promote the interests of farmers. During the post WWI years farmers were short of labor force. The Progressive Party led the government to pursue a two-fold policy on the issue, namely to relocate the unemployed to work in farms ("back to the land" strategy) and to facilitate the immigration of unskilled workers to be incorporated both to the farms and to the industries extractive of natural resources. Due to the specific conditions of the Canadian Winter and the nature of these industries, such strategies created a huge group of seasonal, highly mobile, workers. Depending upon the time of the year, they would be hired in different provinces across the country to perform tasks that demanded nothing but physical effort. As the Depression exposed the weaknesses of this approach, the Conservatives gain office with the promise to end unemployment.

When Bennett took office the demand for relied was increasing across the country and so was the demand for a national, non-contributory, unemployment insurance system. Municipalities, the unemployed, organized labor as well as an important share of Canadian businessmen supported the claim, as illustrated by the fact that in 1932 Bennett was presented with a petition along these lines endorsed by 94169 people. Moreover, the transients issue was growing in importance: according to Whitton’s estimates (1933) there were up to 100,000 seasonal workers in the West on whom unemployment was (and would continue to be) especially concentrated. This large subset of seasonal dependent population, not being residents, highlighted the issue of which layer of government should take care of them. Municipalities and Provinces looked at Ottawa and demanded a response to the added relief burden of the transients. Bennett’s response was rather traditionalist. He insisted on a sharp divide between relief and insurance, putting aside the

---

7 This summary draws on the historical accounts by Guest (1997), Pal (1988) and Struthers (1983). The latter provides the richest historical account of the unemployment problem in Canada between 1914 and 1941.
latter on the grounds of that it would require a constitutional reform of the BNA. He agreed to increase funds to cover the cost of transients relief until July 1933. At the same time he adopted a strict policy of "provincial self-reliance" in what concerned Dominion-Provincial financial relations, constraining the provinces’ capacity to take debt to 15 M a year. Finally, he expanded the "back to the land" strategy by launching the National Defense Relief Camps, a system of concentration camps for "single unskilled men and recent immigrants" in which these would work for the government in exchange for the meals, shelter and money received as relief. The explicit goal was to expel the specter of socialism by drawing transients out of the cities.

These measures proved insufficient. By 1934 the failure of this strategy to guarantee prosperity and the achievements of F.D. Roosevelt’s policies made Bennett change his approach in a two-fold way. He proposed to progressively abandon relief and substitute it with a program of Public Works able to generate employment while at the same time introducing an unemployment insurance system for the bad times to come. Both programs were part of a strategy to overcome the expectable resistance from the Provinces against Ottawa’s withdrawal from unemployment relief. The strategy backfired as it actually alienated the possibility of any such agreement. The proximity of the elections led Bennett to present his own “New Deal” in a series of radio speeches to the nation outlining a number of initiatives to overcome the social and economic crisis of the country. These included a national minimum wage, working-hours legislation, a more progressive tax system and, among others, a contributory old-pensions and health and unemployment insurance. The Liberals highlighted that most of these initiatives would require a reallocation of political capacities from the Provinces to Ottawa but did not oppose the Employment and Social Insurance Act, passed in March 1935. Inspired by the 1911 British System, the proposal for unemployment insurance had stricter benefits, in accordance to the less elegibility doctrine, and very limited coverage (the primary sector was excluded, dealt with only through relief). Moreover, the Government would only cover 1/5 of the total cost of the program and neither sickness nor transitional benefits were considered. In general federal involvement would remain to a minimum in that the Unemployment Insurance would be neither financed nor administered by the Dominion. Bennett’s proposal, remarkably attached to the status quo, was never implemented. The Supreme Court (1936) and the Privy Council in London (1937) declared it ultra-vires, restating the need to amend the BNA if Ottawa was to develop any Unemployment insurance program.

In addition to this major political backlash, the transients issue re-emerged forcefully with the strikes in the relief camps of British Columbia, the organization of the On to Ottawa Trek and the Regina riots. Social unrest kept increasing while the financial relations between the Provinces and the Dominion were broken (the four Western provinces owed Ottawa 117$M.). Not surprisingly, in the 1935 elections Bennett was defeated by a large margin and, with him, "the idea that the care of the unemployed was a local responsibility"(Struthers 1983: 137). The Liberals, led by Mackenzie King, regained office.
King’s moves appeared continuist at first. Indeed, he became tighter on public finances, demanding from provinces a balance budget and aiming at a progressive reduction in the federal share of relief expenditures (mainly via a program of progressive decrease in Grants in aid to the Provinces). However, King also pursued new avenues by appointing a National Employment Commission (NEC) to study alternative courses of action on unemployment. The NEC recommended the closure of the relief concentration camps their substitution for a land settlement program (1936)\(^8\). It also proposed an Employment Service to enhance coordination among employers and employees, an ambitious Public Housing program able to both create employment and lighten the shortages in the housing supply as well as a National Volunteer Conservation System along the lines of the one developed by Roosevelt in the USA (CCC). None but the land settlement program were implemented. In addition, after Bennett’s constitutional fiasco, King also sounded the provinces regarding the amendment of the Constitution, necessary to develop a centralized unemployment insurance program. To this end he appointed in August 1937 the Rowell-Siros Commission on Dominion-Provincial financial relations. At this point six provinces, including British Columbia, agreed to transfer the capacity to deal with unemployment to the federal government. New Brunswick decided to wait for the conclusions of the Commission before making a decision. Quebec and Ontario, provinces with average to high income per capita and relatively less hit by the Depression (figure 6), wary of the fiscal and political implications of a common program, claimed concerns about provincial autonomy and refused the amendment. Incidentally, both provinces are bound to be net contributors in any national insurance system. Finally, Alberta’s Premier (Aberhart), whose earlier attempt to develop a system of Social Credit specifically tailored to the type of production dominant in the province had been overruled by Ottawa, rejected any constitutional amendment unless he could see the specific legislation first. At this point, as a unanimous agreement on constitutional reform was not within reach, King decided to postpone the issue until the Rowell and Siros report was published.

However, exogenous circumstances altered King’s plans. By the end of 1937 a heavy drought increased the magnitude of the Depression in the West, creating yet another conflict between British Columbia and Ottawa about the costs of coping with the transients. While relatively wealthier than any other province, BC was a natural point of destination for the transients. Estimates by the municipal relief office indicate that during 1930 Vancouver’s population increased by 5 percent due to the arrival of 10000 seasonal unemployed begging for care. With its capital dubbed as the “meca of the unemployed”, BC’s financial capacity was strained. In this context, King insisted on a balance budget approach, with proposals to freeze Grants in Aid to the previous year level, to limit up to

---

\(^8\) The conditions under the new program were much better than in Bennett’s concentration camps. Now workers would get a proper salary ("a $5 monthly stipend to both the farmer and person he hired plus a $2.50 deferred bonus for each month worked to be received during the winter months ")\(^\) and the program would be much broader in coverage (targeted at the estimated 100000 transients, it ended up taking care of 37000 homeless men and 5000 women, which were excluded from Bennett’s camps. For a short period of time, this strategy solved temporarily the transients issue and contained the cost of relief expenditures to the provinces. However, British Columbia and Ontario refused to take part in it, limiting its positive effects.
30% Ottawa’s share of total relief costs and to set a maximum level of relief to be provided by the provinces. The proposal exacerbated tensions with the Provinces as well as within his own government. In response, the BC premier successfully forced Ottawa to pay for the relief expenditures generated by all these non-resident workers by closing the camps six weeks earlier and cutting off relief for these workers. Transients had become, again, the most visible symbol of the shared, non-local, nature of the unemployment problem. In December 1937 the NEC issued another report calling for a policy shift towards a national approach combining assistance and insurance. In response to these demands, King’s government put aside the balance budget and agreed to spend $40M on unemployment relief expenditures. However, no Unemployment Insurance Act was proposed as the Rowell-Siros report was not yet concluded.

In this context the beginning of the Second World War added yet another source of cross-regional externalities, ultimately pushing the Government to seek the constitutional reform even before the publication of the Rowell Siros report. The concerns about the future demobilization and the socio-economic integration of veterans, already salient after WWI, rendered unemployment even more national a problem. Taken together, all these factors gradually erased the hesitations among the provinces that either doubted or openly opposed the amendment. Unemployment itself contributed to the replacement of M. Duplessis, Quebec’s conservative premier, by A. Godbout, a Liberal who quickly switched positions. In turn, the level of economic externalities and the War undermined Ontario’s earlier concerns for its fiscal position within the union. Finally, Alberta’s premier, according to Struthers (1983), faced strong pressures as he became the only opponent, on the basis of Alberta’s specific regional economy, to an amendment viewed as a national need for the difficult days ahead. Along these lines, the Rowell-Siros report came at last to back up all the economic arguments in favor of Unemployment Insurance, namely administrative efficiency, inter-provincial equity and especially the fact that all provinces were affected by the problem, i.e., the fact that unemployment "was no longer seen as the result of local conditions". In July 1940 the Unemployment Insurance Act was approved.

A fully centralized program was finally at work. It was heavily based upon Bennett’s legislation, with three important qualifications (Dingledine 1981). First, benefits were no longer at a flat rate, but graded according to previous earnings. Second, a bigger effort was made in terms of coverage (75% of the labor force), eligibility (the qualifying period drops from 40 to 30 weeks) and duration. Nonetheless, seasonal agriculture, forestry and fishing workers remained excluded, treated only under assistance policies still guided by the “less eligibility” principle and with no national minimum. Following the Rowell-Siros report, the full cost of the latter was finally assumed by Ottawa. The report, however, did suggest a centralization of taxation and the design of an inter-provincial system of revenue sharing so that the average provincial cost of relief would end up being equalized.

IV.- The American Response: States retain their freedom to pursue their preferred system of unemployment insurance in the context of a tax-offset system.
The linkage between the social consequences of the Depression and the design of an Unemployment Insurance system is also very clear in the case of the United States (Webber and Wildavsky 1986: 453-464). As in Canada, discussions about Unemployment Insurance date back as far as 1916, when Congress voted to draft legislation on the matter. The proposal did not go forward and the issue died until the 1928-1931 period. As the depth and duration of the Depression were increasing, the Senate Committee on education and labor held hearings "on the national problem of unemployment" (McGowan 1999:3). In 1931 and 1934 Senator Robert Wagner introduced proposals for a federal-states Unemployment Insurance system. Although the bills were not even voted on, the 1934 proposal became a very important precedent of the institutional design finally adopted for unemployment insurance in the United States.

During the 1920s and 1930s the conventional approach to unemployment relief remained unchallenged⁹. Indeed, some features of the traditional view of relief were also present in Roosevelt’s approach to it. In his own words, "the federal government, of course, does have to prevent anyone from starving, but the federal government should not be called upon to exercise the duty until all other agencies fail. The primary duty is that of the locality, the city, county, town. If they should fail and cannot raise enough to meet the needs the next responsibility is on the states and they have to do all they can. It is still proven that they cannot do any more and the funds are still insufficient, it is the duty of the federal government to step in" (New York Times, May 23rd, 1933; cfrd. Singleton 2000:108). The latter was indeed the case. Roosevelt’s action on this matter was much quicker and less reluctant than the one of his Canadian counter parts. Right after taking office he introduced the Federal Relief Emergency Act, "appropriating $500 M to the States" (Singleton 2000:102-110). In so doing he responded "to the plight of the unemployed and to the non inconsiderable protests of state and local welfare officials, whose agencies were overwhelmed financially by the proportion of the need produced by the crisis"(Orloff 1988:69).

As was the case with Alberta in Canada, significant developments in unemployment policy took place at the industry and state levels prior to the Depression. According to the estimates of the Department of Labor of the State of New York, "in 1934 trade unions plans covered about 100,000 workers; joint unions-management plans covered about 65000 and voluntary company plans covered another 70000" (McGowan 1999:5). In addition some states managed to react to the Depression earlier than the federal government and introduced their own unemployment insurance legislations, most prominently the Wisconsin (Nelson 1969) and the Ohio laws (1932). By 1935 six other states (New York, California, Massachusetts, New Hampshire, Utah, Washington) passed their own legislations in anticipation of the Social Security Act. Despite these pioneering efforts, by 1934-35 unemployment insurance was a privilege reduced to a minority of the American labor force. At the same time, the Depression had just reached its peak: the unemployment rate was 25%. A proportional policy response was demanded from the federal government (Weir 1988: 149-199; Skocpol and Ikenberry 1983:87-148). Furthermore the need to face the Depression in a convincing way was urged even more

⁹ Stewart (1930) offers a review of the alternatives available at the time. A detailed historical account of the evolution of unemployment relief during the Depression decades is offered by Singleton (2000).
strongly by the emergence at the left of the political spectrum of a number of much more radical proposals.\textsuperscript{10}

Roosevelt’s efforts on unemployment and old age relief bought him some time to carefully design longer term measures to handle the different dimensions of "economic instability." In 1934 he appointed the Committee on Economic Security to study, among other social security issues, the problem of unemployment. From this point onwards, the history of America’s response to unemployment runs from the production of the report of the Committee through the final approval of the Social Security Act. Along the way, the contentions on the approach and the inputs at different points of the drafting process illuminate why, facing as "national" a problem as Canada, it is finally the states, and not the federal government, that retain full political control of unemployment insurance. The remainder of this section builds on materials produced by direct participants in the process, namely E. Witte’s\textsuperscript{11} The Development of the Social Security Act (1962), Arthur Altmeyer’s\textsuperscript{12} The Formative Years of Social Security (1968) and Paul H. Douglas\textsuperscript{13} Social Security in the United States. An Analysis and Appraisal of the Federal Social Security Act (1936).

In analyzing this material, I first focus on the discussions within the CES. Specifically, I present the different alternatives as to the institutional design of unemployment insurance and discuss the reasons underlying the final choice. Secondly, I analyze the political influences in the re-drafting of the Titles devoted to unemployment insurance of the Social Security Act during its legislative process. This includes the inputs in both the House of Representatives and the Senate as well as the Conference Committee to conciliate the disagreements between both Cameras. Finally, I present a succinct summary of the final outcome.

Douglas (1936:28-69) and Witte (1962: 111-143) concur that there were basically three alternative institutional designs on the table of the Committee on Economic Security. These were: (1) a national system in which the federal government would collect contributions from workers and employers and make transfers directly to the

\textsuperscript{10} Among these proposals were the Long’s plan for a flat universal pension to be given to anyone over 60 or the Townsend's plan to issue a sales tax to pay 200$ a month to every American Citizen except convicted felons (Berkowitz 1991: 1-39). Another example is the Lundeen Bill, submitted to (and defeated in) Congress by Representative Lundeen from Minnesota in 1934 and 1935. The bill provided" for the payment of unemployment compensation out of the funds of the national government to all unemployed persons over the age of eighteen years for as long as they were out of work through no fault of their own"(Douglas 1936: 74).

\textsuperscript{11} Executive Director of the Committee on Economic Security (1934-1935). This book is based on a diary kept by the author during the entire process.

\textsuperscript{12} Second Assistance Secretary of the Department of Labor and Head of the Technical Board of the Committee on Economic Security.

\textsuperscript{13} Professor of Economics at the University of Chicago. He served throughout the years as external assessor as well as discussion leader in the National Conference on Economic Security, a convention held in Washington (1934) where initial discussions on the different types of institutional designs to be implemented were held. Thereafter, he served as close collaborator of Bryce Stewart, Head of the Committee’s Area of Study on Unemployment Insurance and former Director of the Employment Service (1919) in Canada.
unemployed; (2) what Witte calls a "subsidy plan", i.e., a system that "was, in essence, that the amount of revenues collected through the federal tax from employers in each state be returned to that state to be used for unemployment compensation purposes, subject to the state’s compliance with standards to be prescribed by the federal government" (Witte 1962: 115); and finally, (3) a tax-offset system along the lines of the Wager-Lewis Bill (1934), which proposed a payroll tax upon employers equal to 5% of what they pay in wages. If a state passed a law and met the federal standards, "then the contributions paid by employers under such an act would be credited as an offset against the federal tax"(Douglas 1936: 23). Otherwise, states are free to choose whatever system or institutional design for the provision of unemployment insurance. While the the project did not pass through Congress in 1934, the provisions of the Social Security Act follow in the footsteps of the Wagner-Lewis bill.

Despite this outcome, the proposal to adopt a national system of unemployment lacked neither support (the trade unions among others) nor an articulate defense (Douglas 1934: 215-216). On grounds of administrative efficiency, distributive equity, and cost effectiveness, Bryce Stewart and his supporters advocated for treating a "national problem" with a "national solution", following a route similar to the one the Canadian provinces would take in 1940. Indeed, a recommendation for a national system was officially made in October 1st 1934 as part of the first report presented by the Unit of Study on Unemployment to the board of the CES: "if constitutional, a nationally administered system of unemployment insurance is to be preferred to a state system, but the Committee should be satisfied that the nationally administered system is constitutional before commitments in favor of such a system are made to the public".

The paragraph also reflects how the supporters of the tax offset system along the lines of the Wagener-Lewis bill centered their strategy around the issue of the constitutionality of the unemployment insurance system. By that I refer to the need to respect "state rights" while developing policy on those matters, like Social Security, in which the Constitution does not guarantee the Federal Government an exclusive political capacity. Indeed, the Wagner-Lewis bill was itself inspired on the Federal Inheritance Tax Act. Such legislation was the response by the administration of President Coolidge to the attempt by Florida to expand its tax base through the elimination of inheritance taxes. IN response, the federal government introduced an inheritance tax with the provision that "80 % of the sums thus collected would be returned to those states which had state inheritance tax laws. If a state did not have such a law, however, the federal government retained all the amounts paid in from the states residents" (Douglas 1936: 22). Florida presented a case for the autonomy of the states before the Supreme Court and lost it. Aware of this precedent, supporters of the tax offset system were closer to Roosevelt’s own views. In the same Declaration of Policy (June 8th 1934) where he announced the appointment of CES, he indicated that "he favored a plan providing for a maximum of cooperation between the states and the federal government, leaving to the states a portion of the cost of management and to the federal government the responsibility of investing, maintaining and safeguarding the funds constituting the necessary insurance reserves". 
While Roosevelt’s administration was concerned with Social Insurance being potentially turned down by the Supreme Court (Douglas 1936:33), this fear did not stop him in the case of old-age insurance. Thus, I would argue, what really resonates behind his words are strictly political concerns about the potential opposition of the Southern States to any system disrupting their specific system of labor relations. Roosevelt effectively encouraged the CES to factor in any issue that could jeopardize the passing of the Social Security Act (Noble 1997; Berkowitz 1991). Francis Perkins, Labor Secretary, and Witte took this encouragement as a mandate, and made every effort to prevent the final CES report from recommending a national unemployment insurance system.

In addition, the early risers of unemployment insurance at the state level worried about a national program. They saw it as a political risk in that it would eliminate provisions tailored specifically towards state’s labor markets and impose dysfunctional policy designs. Witte himself, a member of the Commons School instrumental in the development of the Wisconsin model of unemployment insurance (Nelson 1969), exemplified this position, and so did Wagner from New York (Altmeyer 1968: 18-24; Orloff 1988: 69-75). Combined with Roosevelt’s recommendations, these reservations shaped deliberations within the CES. On November 9th a motion according to which "all the thought of an exclusively federal system be abandoned" (Witte 1962:118) received approval. By November 14th Roosevelt announced that unemployment insurance would be a "federal-state undertaking". In other words, there would no national unemployment insurance system in the United States. The question had been narrowed down to a choice between the "subsidy plan" and the tax offset system. After several rounds of consultation between CES and the Advisory Committee of the Department of Labor (initially in favor of the subsidy plan for 9 to 7 votes), it was decided to draft the Unemployment insurance provisions along the lines of the Wagner-Lewis Bill, not only because it guaranteed to each state the possibility of pursuing its own strategy, but also for reasons of constitutional pragmatism. According to Douglas (1936: 48), "the one major argument that came to be urged for the offset plan as opposed to the tax remission system was that, if the latter were later to be declared unconstitutional, the whole system of unemployment insurance would necessarily collapse, whereas, even if the tax offset method were finally rejected by the Supreme Court, the states would in the meantime have passed acts that would continue".

In August 18th 1935 Roosevelt signed the Social Security Act, with the following provisions on unemployment insurance:

1. Unemployment insurance is a joint effort between the federal government and the states, and takes the form of a tax offset system in which the participation of the states is voluntary.

2. The federal government levies a payroll tax on employment that would be equal to 1% of total payroll in 1936, 2% in 1937 and finally equal to 3% from 1938 onwards. This tax is deposited in a federal unemployment trust fund under the control of the Secretary of the Treasury. 90% of the revenues are devoted to transfers to the states that effectively
develop their own unemployment insurance system. The remaining 10% is devoted to cover administrative costs.

3.- States are free to choose the specific institutional form of their unemployment insurance system (plant reserves, industry reserves etc.). They are also left free to adopt any scale of benefits they wish. Furthermore they have full control on waiting periods and duration of benefits.

4.- State’s Unemployment insurance systems must none the less conform to a number of standards established in the law, most of which come directly from the Wagner- Lewis Bill. These are:

4.1.- The payment of benefits must be made through the public employment offices of the state or in any other form "approved by the Social Security Board".

4.2.- No state system can disqualify a recipient if he refuses to take up a job vacant due to any kind of labor dispute.

4.3.- The same applies to the unemployed who refuse work which pays "substantially less" than the ongoing wage rate and/or has "substantially worse working conditions" attached to it. The specific meaning of "substantially less" and "substantially worse" is left to be decided by the states within a range (or zone of tolerance) established by the Social Security Board.

4.4.- The same applies if the worker refuses to sign a "yellow-dog" contract. A "yellow dog" contract is one that requires the worker to join the company union and/or refrain from joining "bona fide labor organizations" (Douglas 1936: 137).

5.- Finally, the Social Security Act excludes from the unemployment insurance system a number of specific occupational categories and businesses. These are: agriculture (including croppers and tenants), domestic service, shipping, public employees (federal, state and local), non-profitable organizations and the self-employed. In fact every firm with 7 or less workers is exempted from paying the federal payroll tax.

Interestingly enough, there is a significant distance between these and the provisions of the original bill as drafted by the Committee on Economic Security. A very illustrative political process took place in the interim, where conflicts about the details concentrated the bulk of political contention, thus shedding light on the ultimate causes behind the selection of a decentralized unemployment insurance system. These conflicts were to take place in the Ways and Means Committee of the House of Representatives as well as in the Finance Committee in the Senate. Table 1 presents the diachronic sequence of the legislative process by comparing its four main stages, namely the original draft, the changes introduced in the House and the Senate and, finally, the agreements achieved.
in the Conference Committee. The comparison is structured around four aspects of the law: the revenues to pay the unemployment funds, the institutional design of the states systems, the coverage of the program and the provisions for the administration of the system. Two aspects are particularly noteworthy: the La Follete amendment in the Senate and the evolution of provisions on coverage.

[TABLE 1 about here]

The original version of the bill granted states freedom to design the provision of unemployment insurance. However, reforms introduced in the House limited the number of designs the unemployment insurance program could take on grounds of accounting efficiency. This meant, among other things, that states were no longer permitted to adopt or maintain systems of industry or plant reserves. Should this provision have passed, the Wisconsin system had to be reformed. As one of the pioneers of unemployment insurance in the United States, Wisconsin fought successfully to protect the system they thought best served their economy. Senator Robert Lafollete Jr., the junior senator from Wisconsin and member of the Progressive Party, successfully introduce an amendment to restore the original provision. Paradoxically, to protect the early developments in unemployment insurance, he facilitated the task of those who wanted to ensure as late and little a policy as possible. Southern states were also successful in limiting federal control over the implementation logistics. They succeeded in abolishing the requirements of meeting the Federal Standards of Personnel and organizing payments only via public employment offices (see the section on Administration in Table 1).

The introduction of specific provisions excluding agricultural workers of the system of unemployment insurance highlights even more clearly the weight of the South as the political engine underneath the process. Alston and Ferrie (1999: 49-75) argue that Roosevelt and CES were aware of this and managed to anticipate potential objections when they decided (1) to exclude firms with less than 4 workers and (2) re-classify croppers and tenants as agricultural workers. They thought both measures would be enough to keep the system of labor relations in the South unaffected by the new unemployment insurance system. Southern Democrats wanted more guarantees (Brinkely 1984; Whatley 1983: 905-930): hence the specific changes, introduced both in the House and the Senate, on the threshold of exclusion on the one hand and, more directly, on the occupational categories left out of the program. After these changes, the Social Security Act constituted no threat.

---

14 The Conference Committee is a joint Committee where conferees from both Houses meet in order to reach a consensus on those specific points of the legislation where the decisions of the House and the Senate have previously differed.

15 An even more straightforward example is offered by the changed in the legislation about old-age insurance. Southern Democrats managed to erase from the Social Security Act any paragraph that potentially could lead to the inclusion of "negroes" in the program. So for instance they erased from the SSA an special Old Age insurance program recommended by CES. They also eliminated a general requirement in the law according to which "assistance could not be denied to any US citizen". However, during the post WWII period, old-age social security was much faster than unemployment insurance in abandoning these restrictions. For instance, the inclusion of agricultural and domestic workers did not
In conclusion, and to put it in terms of the analytical model, the American response includes an increase, through tax-offsets, of transfers to regional governments (T, in the analytical model above) so that each state can design at will its own system of unemployment insurance (that is to say, a decentralized t above). In turn, Canadian political elites for a fully centralized system (FC).

IV.- Comparative Analysis: Risk-sharing and the Centralization of Social Insurance.

The very existence of two different outcomes poses a puzzle for single factor explanations based either on the existence of institutional veto points or on the need to pool resources to face a structural economic shock. The American appears consistent with the former as it has been shown by a large number of contributions that explain why the United States is a welfare laggard. But the Canadian experience points to the existence of other factors that may lead the units of a federation to alter the institutional design and renounce, through an amendment, their constitutional veto capacity on a national program for unemployment insurance. On the other hand, given the magnitude of the Depression, one could see the Canadian response as the natural one: such a shock must necessarily lead to pooling resources. Then why did the United States take a different route?

According to the model developed in this paper, the scope of risk sharing between regions determines the institutional design of social security. More precisely, the theoretical model points to three potential keys to understand the divergent paths between Canada and the United States: income inequality between regions, the size of the dependent population by region ($\frac{\alpha}{1-\alpha}$), the extent to which the level of dependents mobility homogenizes this ratio across regions, and the incidence of risks associated with specialized economic activities ($\delta$). Similarly uneven in both countries (see Figure 6), the geography of inequality cannot be the ultimate reason for the institutional divide in response to the Great Depression. In terms of income, Canadian wealthier provinces had incentives to prevent a centralized system of unemployment insurance. Yet they did not. In turn, some of the poorest American states worked, successfully, to prevent a national program that, in pure income terms would benefit them. If the argument of the paper carries any leverage, the key must lie elsewhere, namely in the different balance between the scope of interregional economic externalities, primarily the level of dependent’s mobility, and the extent to which specific regional economies shape the geography of preferences about public insurance systems. The remainder of this section evaluates this claim in detail.

---

16 See for all Quadagno (1994) and Noble (1997). Incidentally, the consideration of the USA as a welfare laggard must be treated with caution. This label is only applicable to the second half of the twentieth century. Lindert (1994) offers estimates of the extent of redistributive social spending for 1910 and 1930, concluding that Continental Europe does not outperform the United States in the First Half of the twentieth century. See also Skocpol (1992).
During the late XIX and early XX centuries, racism shaped the "unfinished democracy" (Quadagno 1994) at work in the Southern states. The Ku Klux Klan was spreading into the plantation areas and many black families found it rational to migrate to the urban areas either in the North or in the West (Steckel 1983). According to historical census data, the proportion of blacks born in their state of residence fell from 83% in 1910 to 75% in 1930. In this context, concerns about an adequate supply of labor grew among planters. According to Alston and Ferrie (1999: 17) "some planters chose a new course-turning to honesty, fair dealing and a host of nonwage aspects of their relationship with their workers as additional margins for competition". Over time a system of mutual obligations emerged and consolidated. "By the early twentieth century planters had come to act as intermediaries between their workers and much of the outside world. Planters exercised control over the credit extended to their workers, but they were also willing to stand good for their worker’s debts with local merchants. [...] Planters reported significant outlays for the payment of doctor’s bills, the establishment and maintenance of schools and churches, and various unspecified forms of entertainment. And planters commonly paid legal fines incurred by workers and served as parole sponsors for their workers" (Alston and Ferrie 1999: 20). In return, workers, especially tenants and croppers, "were expected not only to work hard in the fields but to display deference toward their landlords"(ibid: 25). Such basic exchange underlies what economic historians conventionally refer to as Southern Paternalism.

Its logic was compelling. Tenants and croppers gained protection and side benefits for their families. From the planters point of view an adequate supply of low-cost labor was guaranteed. The longer the duration of the exchange the bigger the levels of mutual dependency and hence the larger the opportunity cost of moving away for croppers and tenants. Put differently paternalism reduced mobility and hence maintained the specificity of the planters’ system of production and social dominance. In fact, the proportion of blacks born in their state of residence did remain constant between the 1930 and the 1940 censuses. These in turn had incentives to keep "looking after " their tenants as well as to prevent any external input potentially disruptive of this particular relation.

The key for the system to work was the mutual dependency between planters and tenants/croppers. Hence when, in the context of the Depression, several federal relief and insurance programs (including unemployment insurance) were under consideration, planters faced the following dilemma: " How could it accept the government assistance that so many plantation owners desperately needed to sustain their labor force until prosperity’s return without allowing the government to replace them as the benefactors of their workers?" (Alston and Ferrie 1999: 49). Direct federal provision of relief, FERA, (Williams 1939) and insurance in favor of croppers and tenants would have undermined their dependence upon the planters and, eventually, led to the ultimate crisis of paternalism: hence the trade-off between federal welfare and the protection of a specific labor market. And hence the final choice of the Southern elites: to press for the preservation of "states rights" in the management of welfare provision as well as, more specifically, for the successful exclusion of their workers from any type of insurance program, including old-age and unemployment. Given the institutional position of
Southern representatives in the House and the Senate, "agricultural labor and domestic services were excluded [from unemployment insurance] as a matter of course" (Witte 1962: 132). In addition, other states also joined the South in a defense of regional specific interests, even if of a different kind. Such is the case of the pioneers of unemployment insurance in Wisconsin for instance, as proved by Lafoyette’s intervention during the legislative process in defense of the states’ freedom to tailor unemployment insurance programs to the specific needs of regional economies. Overall, differential risk profiles across states played a crucial role in shaping the American response to the Depression.

The structure of incentives was different for the Canadian provinces. Although there were different degrees of specialization among them, and some Premiers like Alberta’s defended them throughout the process, there was nothing alike the Southern American states in terms of the specificity of production and the peculiar internal equilibrium of the system of labor, social and political relations. In brief, even in the context of the Depression, the labor market specificities of the American States (including not only the Southern ones, but also some Northern and Midwestern ones) reduced the levels of risk sharing and hence made decentralization the most likely institutional design for unemployment insurance. In terms of the model, inter-state differences in the value of $\delta$ were higher in the United States than it was in Canada, thereby creating political incentives to adopt a more decentralized insurance system.

The differences between the two North-American federations go beyond the geography of labor market risks as Canada and the United States also show distinctive patterns of dependents mobility. The historical account above shows how, after WWI, Canada and the United States took opposite paths on the issue of immigration. While the United States increasingly made a more selective policy, Canadian borders were open to large amounts of unskilled labor, let in to work primarily in farming and extractive industries. These immigrants were from the beginning a very mobile sector of the Canadian labor force in that their familiar and cultural ties did not belong to any particular province. Thus, when the Depression affected both countries, the transients were a Canadian peculiarity, a mass of seasonal workers moving between jobs and the quest for welfare, thereby functioning as a multiplier across territories of the social consequences of the Depression.

By moving both eastwards and westwards, transients passed the effects of all these adverse shocks to the more industrialized provinces of the country. For instance, according to Struthers (1983:44-71), transients from the West constituted 40% of the total demand for relief only in the city of Toronto. As mentioned earlier, Vancouver suffered a similar shock by receiving over 10000 people in need of support by the end of 1930. Estimated to be over 100000, the transients worked to homogenize risk profiles across Canadian provinces. To put these numbers in context, the transients amount to 3% of Ontario’s total population in 1931. More importantly, between 1931 and 1941, Alberta and Saskatchewan suffered a net immigration loss of, respectively, 41841 and 157545 people, amounting to 5 and 18 percent of their population at the outset of the Depression. During the same period, Ontario and British Columbia experienced a net immigration
gain of, respectively, 77484 and 82498 people. These population moves made the incidence of the Depression to be territorially much more balanced.

In terms of the model provinces became much closer in their value of \( \frac{\alpha}{1-\alpha} \), thereby homogenizing the demand for welfare and insurance across territories. Likewise, transients also contributed to a reduction in the gap between provinces in the value of \( \delta \): being mobile across provinces between seasons, transients and their associated risks became a shared sector of the labor force. Put shortly, transients did contribute to the increase of the levels of risk sharing between provinces and, thus, to making centralization the preferred institutional design for unemployment insurance. Meanwhile, despite the Dust Bowl exodus immortalized by John Steinbeck, no such group of seasonally mobile rural workers/dependents was formed in the United States as Southern paternalism worked purposefully to limit the inter-state mobility of agricultural workers.

The centrality of interterritorial mobility to understand the structure of incentives underlying different institutional designs is reinforced by the evolution over time of American social policy. When as a consequence of the WWII the Western American states demanded more labor supply, the Southern states managed to pass the Bracero program (1942-1964), a policy that arranged the immigration of temporary Mexican laborers for that purpose. Thereby they succeeded once again in preventing an external influence to undermine the dependence of their croppers and tenants (Alston and Ferrie 1999: 99-119). Over the years, Southern paternalism would disappear because of technical change and mechanization. Only then African-Americans became more mobile entering the national political economy, gaining civil and political rights. Indeed, the rate of African-Americans born in their state of residence started to decline again between 1940 and 1950, going from 76 to 68 percent. Focusing specifically on the South, Quadagno (1994 reports that while in 1940 77% of African Americans lived in the South, by 1970 only 53% still did. As a result, “the presence of lack migrants in northern cities moved racial inequality from the periphery to the center of national politics” (p.25). Again mobility increased risk sharing, eventually contributing to a major step forward in the development of the American welfare state, Lindon Johnson’s War on Poverty and Equal Opportunity Act (1964). Only then, “the barriers to equality of opportunity the New Deal had created” were challenged (p. 31).

V.- Conclusion

The comparative historical analysis of the emergence of unemployment insurance in Canada and the United States offers a good deal of empirical support to the argument developed in this paper. Canada adopted a centralized system in 1941 because the scope of the interprovincial externalities was strong enough to overcome provincial concerns. The income base of wealthier provinces, such as Ontario and British Columbia, was effectively undermined already by the inflow of dependents, whereas the ability to resist by those relatively poorer members with alternative designs tailored to their provincial needs, such as Alberta’s premier, was undermined by their overwhelming income needs. The Canadian experience illustrates how, as suggested by the model, large scale

\[17 \text{ Source: Statistics Canada. Historical Statistics (online version).} \]
interregional externalities can overcome the fragmenting effects of an uneven economic geography, even under conditions of centrifugal representation. The balance between regional specificities and interstate externalities was reversed in the USA. Regional economic specificities were much stronger, particularly in the race based political economy of the South, whereas mobility data suggest that interstate economic externalities, while present, were not as pressing as in Canada. Under these circumstances, the fragmenting effects of the geography of labor market risks constrained the development of a centralized system of unemployment insurance. In turn, consisted with the model, the USA launched a system that combined moderate levels of horizontal redistribution via a tax-offset system with the protection of a decentralized provision of unemployment insurance. In sum, given a system of representation that privileges regions, the balance between interregional economic externalities and the geography of inequality and labor market risks drives the design of public insurance systems.

The argument and findings of this paper relate to existing scholarship in a number of ways. Contrary to the dominant notion that federalism and decentralization are necessarily inequalitarian, this paper has shown how similarly decentralized political systems can vary in the organization of their fiscal structures, as well as in the scope of their efforts to prevent the spread of inequality (Lindert 2004; Obinger, Leibfried et al. 2006). In addition, it has offered a theory as to why these variations occur. The argument brings insights from the literatures on economic geography (Krugman 1991; Venables 2001) and factor mobility (Wildasin 1991, 1995; Cai and Treismann 2005) into the literature on endogenous fiscal institutions (Persson and Tabellini 2000; Bolton and Roland 1997; Alesina and Spolaore 2003; Beramendi 2007; Wibbels 2005). In doing so, this paper identifies risk-sharing between regions as a key factor, separate from income, shaping the political economy of fiscal institutions. More specifically, the paper highlights the mobility of dependents as a particularly relevant factor behind crossregional externalities, ultimately accounting for the design of public insurance programs. Herein lies an important key to understand why some federations have more centralized welfare states than others, and ultimately why some federations redistribute more than others. In making this case, the paper moves beyond the voluminous literature on mobility and redistribution in federations to place mobility and its interplay with the geography of labor market risks at the center of a theory of institutional design. Finally, the argument also contributes to illuminate the observable associations between fiscal institutions and distributive outcomes. Exploring the endogenous nature of public insurance systems improves our understanding as to why the United States remain an outlier in terms of redistribution and inequality among advanced industrial societies.

A number of limitations demand further attention. Chief among them are the need to expand the analysis to other systems of representation (Persson and Tabellini 1996a, 1996b; Beramendi 2008), the issue of separating and better identifying the joint endogeneity between mobility and institutional changes, and the need to expand the array of policies under consideration, as the strategic interaction among the relevant actors may be subject to policy-specific constraints.
REFERENCES


Card, David. and Freeman, Richard. 1996. Small Differences that Matter, NBER.


<table>
<thead>
<tr>
<th>ORIGINAL BILL</th>
<th>HOUSE OF REPRESENTATIVES (Ways and Means Committee)</th>
<th>SENATE (Finance Committee)</th>
<th>CONFERENCE COMMITTEE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>REVENUES (Federal Payroll Taxes)</strong></td>
<td>The rate of the federal payroll tax is related to previous years' levels of production. The normal level is thought to be 3%.</td>
<td>The rate is to be established following an annual scale. 1935: 1%; 1936: 2%; 1937: 3%. Thereafter, it ceases to be related to previous years' levels of production.</td>
<td>Remains</td>
</tr>
<tr>
<td><strong>INSTITUTIONAL DESIGN</strong></td>
<td>Freedom to adopt any type of unemployment insurance system.</td>
<td>The options are limited on grounds of accounting efficiency. States are no longer permitted to adopt systems of industry or plant reserves. Plant reserves were also said to be too inequitable.</td>
<td>La Follette amendment passes. Previous freedom is restored</td>
</tr>
<tr>
<td><strong>COVERAGE</strong></td>
<td>Wide coverage. No specific exclusions beyond public agencies. Only those firms with less than 4 workers are exempted from the federal payroll tax. Specific occupations (mainly agricultural workers, including croppers and tenants) are excluded. The threshold of exemption is increased up to firms with 10 workers.</td>
<td>Remains</td>
<td>The threshold of exemption is set back to 4 workers.</td>
</tr>
<tr>
<td><strong>ADMINISTRATION</strong></td>
<td>Federal Standards of Personnel are required to all States. Benefits are to be paid only by the Public Employment offices of the States.</td>
<td>Abolishment of Federal Standards of Personnel.</td>
<td>Remains</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sources: Data collected from Douglas (1936), Witte (1962) and Altmeyer (19)

---

13 If the index of production of the Federal Reserve Board was by September 1935 less than 84% of its 1923-25 average, the payroll rate would be 1%. If the index was between 84 and 95%, the rate would be 2%. If the index was above 95% of the 1923-25 average, the rate would be 3%.

19 In the case of the states willing to go for an industry or plant systems, it was provided that they should contribute 1% of their payroll to a pooled state fund in order to generate a reserve for other industries.

19 Note that under these conditions the existing Wisconsin system would be doomed to change or disappear.