University of Heidelberg

Department of Economics

Discussion Paper Series No. 441

Rational ignorance is not bliss: When do lazy voters learn from decentralised policy experiments?

Jan Schnellenbach

May 2007

Rational ignorance is not bliss: When do lazy voters learn from decentralised policy experiments?*

(this version: May 30, 2007)

Jan Schnellenbach

Ruprecht-Karls-Universität Heidelberg – Alfred-Weber-Institut – Lehrstuhl für Finanzwissenschaft – Grabengasse 14 – 69117 Heidelberg – Germany – jan.schnellenbach@awi.uni-heidelberg.de

Abstract.
A popular argument about economic policy under uncertainty states that decentralisation offers the possibility to learn from local or regional policy experiments. We argue that such learning processes are not trivial and do not occur frictionlessly: Voters have an inherent tendency to retain a given stock of policy-related knowledge which was costly to accumulate, so that yardstick competition is improbable to function well particularly for complex issues if representatives' actions are tightly controlled by the electorate. Decentralisation provides improved learning processes compared to unitary systems, but the results we can expect are far from the ideal mechanisms of producing and utilising knowledge often described in the literature.

JEL Classification: H73, O31, D83
Keywords: Policy decentralisation; fiscal competition; model uncertainty; collective learning.

*Earlier versions of this paper have been presented at seminars in Berlin, Heidelberg, Marburg and Silvaplana, and at annual conferences of the Royal Economic Society, the International Institute of Public Finance, the Verein für Socialpolitik and the Association for Public Economic Theory. I would like to thank the participants there for comments, and in particular Lars P. Feld, Arye L. Hillman, Federico Revelli, Alois Stutzer and Heinrich W. Ursprung. The usual disclaimer does, of course, apply.

1. Introduction

In one of his most often-cited papers, Hayek (1968) argues that competition on the marketplace serves as a "discovery procedure" for new, previously unknown problem-solving routines. This argument, which is probably more or less undisputed with regard to the market for private goods – see, for instance, Kerber and Saam (2001) and the literature cited there – has recently also been proposed with regard to public goods: Competition between jurisdictions is supposed to provide incentives both to conduct experiments with new policy routines and to learn from experiments conducted in other jurisdictions. This basic argument has, for example, been made by Salmon (1987), Vihanto (1992) as well as Vanberg and Kerber (1994). In spite of growing interest, the observation made by Oates (1999), that this is still a relatively little researched problem, appears to remain valid to date. There are a number of contributions concerned with the general question if political competition induces representatives to experiment with novel policies (e.g. Rose-Ackermann 1980, Kollman et al. 2000, Strumpf 2002). While these contributions come to conflicting results regarding representatives' reactions to political competition, they all tend to put emphasis on the decision-making of politicians, while our focus in the present paper is on the question: Do parallel political experiments in multiple jurisdictions lead to better informed voters?

The above-mentioned contributions have another, not always explicitly made, assumption in common, namely the assumption that individuals have some motivation to gather information on institutional evolution or on the evolution of economic policy in other jurisdictions, and to update their knowledge accordingly. In the contribution by Vanberg and Kerber (1994), for instance, this motivation is explained by referring to the private gains that can be accrued from efficiency-enhancing institutional change. This is certainly very plausible when we are interested in individuals considering their "exit"-option: If an individual considers herself mobile and has the option of leaving jurisdiction A for jurisdiction B, and if she can gather information about B at very low cost, then she obviously has an incentive to inform herself about the real disposable income that she can earn in B. This incentive

disappears, however, when the option of mobility does not exist and the only remaining option is "voice". In this case and from the perspective of a single voter among many, a change of policy is a pure, Samuelsonian public good. If there is not a sufficiently high probability for a single voter to cast the decisive vote, and if there are no external rewards for her changing her mind, then the individual has no obvious incentive to burden the costs to update her given, individual stock of economic policy-related knowledge. She is rationally ignorant in a Downsian sense and has, behind a veil of insignificance, no incentive to invest resources into holding the scientifically most accurate point of view. This problem, however, seems to be widely neglected in most of the available literature on political competition.

This confronts us with a seemingly paradoxical theoretical question: How do rationally ignorant voters learn, who in a strict sense do not have any incentives to learn at all? At this point, we do not want to engage in a fundamental discussion regarding the appropriateness of Bayesian learning as a modeling technique in general (for a thorough discussion of that, see e.g. Hodgson 1997 and the literature cited there). But we believe the conclusion to be straightforward that it is not the appropriate choice under the particular incentive structure sketched above. The costs involved with Bayesian updating, be it in terms of cognitive costs in information processing, or in terms of the costs of gathering information, appear to be prohibitively high given that the private gains of this procedure are essentially nil here. We therefore decide to model learning as a collective, social communication process which requires next to no effort at all on behalf of the single voter. Arguing from there, it will be shown that important implications for the theoretical concept of decentralised economic policy-making as a discovery procedure arise. Under plausible parameter configurations, learning from decentralised political experiments may not occur at all, or, which is a result that is new to the literature, political experimentation may occur only in the wrong (the relatively efficient) jurisdiction. Nevertheless, it is argued that even under such unfavourable conditions, decentralised policy still offers more scope for expanding the stock of knowledge in the long run than centralised policy.

To be able to focus on collective learning processes involving voters and citizens, we assume a tightly controlled government throughout the paper. The policies preferred by the majority of voters are executed frictionlessly and there are no control problems to be solved. In the present framework with its emphasis on uncertainty regarding the true model of the economy, an alternative approach could assume that, like their counterparts in private firms, politicians can engage in cognitive leadership (Witt 1998). This would be a plausible extension of the model developed in this paper, but to keep the analysis tractable, we will restrict ourselves to allow only collective learning of voters at this time.[1] The argument will proceed as follows: In the following section, the dissipation of policy-related theories within a population will be modelled as a frequency-dependent process leading to a stable equilibrium with a clear-cut majority theory. *Section 3* introduces a hypothetical, yet very general starting point for factor migration as well as a distinction between loyal and perfectly mobile individuals. *Section 4* discusses the incentives following from factor migration to critically examine given policy routines and to experiment with new routines. Finally, *Section 5* offers some conclusions.

2. Individual uncertainty and the emergence of common beliefs

2.1. **The dissemination of policy-related conjectures.** The point of departure of the argument presented here is fundamentally different from that of approaches to fiscal competition which involve omniscient maximisers of welfare, rents or something alike and ask whether such a maximising effort by a number of decentralised social planners leads to a result that would be considered optimal by an omniscient, centralised social planner. Instead of following this lead, the notion of model uncertainty is used here: individuals are theoretically uncertain in the sense that they do not know the true model describing the actual properties of the economy within which they are acting and making decisions.

[1] A discussion of the normative aspects of public entrepreneurship can be found in Schnellenbach 2007.

Because the quality of economic policy is a public good and because we assume that there is a large number of citizens, so that the individual probability of having the decisive vote is approximately zero, individuals do not feel a need to invest into acquiring "rational expectations" regarding economic policy, i.e. to utilise all available information in order to gain the most precise theoretical and empirical knowledge about their economy that can be gained at a given point in time. Individuals might be expected to build rational expectations if the necessary information was available costlessly and if it could be learned effortlessly. But both requirements are not met here.

Following Hirschman (1989), however, it is assumed that individuals do feel an intrinsic need to have *some* point of view on issues of economic policy – but, given the public good problem, they do not feel a need to take the scientifically most up to date point of view. On the contrary, it is assumed that, once individuals have learned a set of conjectures about different economic policy measures, they will attempt to retain them. To explain this tendency, assume that at a time $t = 0$, a representative individual is completely uncertain and has no a priori knowledge at all to fortify an opinion on economic policy. Given her assumed intrinsic need for such an explanation, she will acquire some theory $\Omega^n \in \{\Omega^1, \ldots, \Omega^N\}$ that is supplied to her in the public discourse. The supply side of the theory market is not explicitly modelled here. Following Lord Keynes' famous quote that "*Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist*" (Keynes 1936, chapter 24), one may simply assume that every Ω^n has been introduced by economists into the public discourse. This admittedly arbitrary assumption is of course unsatisfactory from the point of view that a theory on economic evolution ought to explain not only the relevant communication (replication) and selection mechanisms, but also the emergence of variety, in our cases a variety of theories, in the first place (see Hodgson and Knudsen 2006). However, treating this stage explicitly would render the argument presented in this paper much more complex, and at the same time add little to understanding the core question of learning from political experiments. Thus, we believe that the simplifying assumption is warranted here.

As a preliminary to explaining the choice of an Ω^n by an arbitrary individual, we assume for simplicity that citizens are homogeneous with regard to their maximand. They all wish to maximise the same objective of economic policy, which in our case is the level of disposable income, but it might just as well be employment, output growth rates or something alike. Since we focus on the general learning process about economic policy measures and not on some specific, well-defined policy problem, we do not need to concern ourselves with the details of the maximisation problem here and can simply assume that there is a common maximisation problem which concerns economic policy-makers and citizens.

In this case, a plausible criterion for choosing one Ω^n among a possibly large number $N > 0$ of available sets is the number of individuals who are already convinced that Ω^n gives an accurate description of the true working properties of the economy. If one is completely uncertain about the relative accuracy of the $N > 0$ available theories, then the number of individuals who already hold an Ω^n may be interpreted as a signal for its usefulness relative to the other theories. It also may be the case that the uncertain citizen decides upon choosing an Ω^n following personal communication with other, already decided individuals. In this case, the probability that the uncertain individual communicates with an individual advocating Ω^n will usually rise with the fraction of already decided individuals who adhere to that set of conjectures.

Thus, it should be possible to model the individual selection of a set of conjectures about economic policy as a frequency-dependent process:[2] A relatively large number of individuals who already hold an Ω^n reassures an uncertain individual that Ω^n is not an obscure, but a reasonable choice. One tool among others to model such processes of frequency-dependent self-organisation is the generalised Polya process,

[2]To some degree, there is obviously a similarity to Kuran (1987) here, in the sense that individuals decide on taking a certain position according to the number of other individuals who already hold that position. The important difference, however, is that in our model there is no place for preference falsification: There is no difference between what individuals privately believe in and what they publicly advocate. Nevertheless, the result, for which Kuran coined the term "collective conservativism", will be quite similar.

as proposed by Arthur, Ermoliev and Kaniovski (1983, 1987). The essence of this process is shown in (1),

$$E[w_{t+1}^n|w_t^n] = w_t^n + \frac{1}{m+t}\big(q_t^n(w_t^n) - w_t^n\big) \quad \text{with} \quad \sum^N q_t^n(w_t^n) = 1 \tag{1}$$

which simply states that the expected value of the fraction $0 \leq w_{t+1}^n \leq 1$ of individuals in the population of already decided individuals who adhere to an Ω^n at a time $t+1$, given its fraction at a time t, depends primarily on just that w_t^n and on an arbitrary, upward-sloping function $q_t^n(w_t^n)$. Time in this model is equal to the number of individuals who have decided themselves, i.e., it is assumed that at any point in time exactly one individual decides which theory to choose. The parameter m stands for the number of individuals who were already decided at $t = 0$ and henceforth, we will simply assume $m = N$, with the underlying assumption that every Ω is backed by exactly one individual at $t = 0$. From (1), the condition for an equilibrium is easily inferred. Beyond identifying the necessary condition for an equilibrium, existence has been proved in the original work by Arthur, Ermoliev and Kaniovski.

Lemma 1. *With q_t^n being defined as a non-negative and non-decreasing function of w_t^n, there must exist at least one fixed point with $E[w_{t+1}^n|w_t^n] - w_t^n = 0$, with each fixed point satisfying $q_t^n(w_t^n) - w_t^n = 0$.*

Proof. See Arthur, Ermoliev and Kaniovski (1983). □

2.2. **Choice and equilibria on a theory market with heterogeneous individuals.** The piece that is still missing in our depiction of the market for theories on economic policy is a set of assumptions on the shape of the function $q_t^n(w_t^n)$, assigning a probability for the next uncertain individual to choose Ω^n to the current market share of this theory, w_t^n. If individuals at any point in time deterministically chose that Ω^n with the highest current market share, matters would be rather simple: The first individual at $t = 0$ would choose randomly one theory to become the most-frequented one, and unconstrained herding behaviour would lead all

subsequently deciding individuals to choose exactly the same. The process would be locked in on a path towards a stable equilibrium with $w_t^n = 1$ for the majority theory, immediately after the first individual has made her random decision. Obviously, the resulting complete consensus among individuals regarding their beliefs about the proper economic policy contradicts even casual empirical evidence.

As an alternative, consider the situation when individuals are heterogeneous regarding their tendency to follow the majority. Let α denote the individual tendency to be conformist, with an $\alpha \in (-\infty, 0]$ signifying a strictly conformist individual who does always and uncompromisingly choose the majority opinion and an $\alpha \in [1, \infty)$ signifiying a strictly nonconformist individual who always and uncompromisingly refuses to take the majority opinion. Values of $\alpha \in (0, 1)$ reflect different degrees of conformism, with the actual choice depending on w. For example, an individual with a relatively high value of α just below unity is a relatively non-conformist individual by nature, but a very high w_t^n may still convince her to join the majority. Nonconformists are thus not completely deterred by large groups with an internal consensus – they just need a relatively bigger consensus group to convince them. When 6 out of 10 individuals report to a non-conformist that they have seen a black swan, he might remain unconvinced, but when 99 out of 100 report the same, even he is likely to believe in the existence of a black swan in the neighborhood. In other words, being a non-conformist is not the same as being intellectual hermit. On the contrary, a non-conformist is an individual who has a tendency to oppose the majority, but not an individual who seeks to distinguish himself by holding a deliberately obscure point of view. Even people who have a strong enough nonconformist tendency to pick the minority theory have a preference to be in a larger minority group, rather than a smaller minority group.

Let Ω^* denote the most popular theory at any given time,

$$\Omega_t^* = \arg \max_{\Omega \in \{\Omega^1, \ldots, \Omega^N\}} w_t^n(\Omega^n). \tag{2}$$

If there is no unique Ω_t^*, but a set of equally popular theories, then Ω_t^* is chosen randomly from this set, with equal probabilities of choice attached to each equally popular theory. Assume individuals value the available theories according to (3):

$$v(\Omega) = \begin{cases} (1-\alpha) \cdot w(\Omega) & \text{if} \quad \Omega = \Omega_t^* \\ \alpha \cdot w(\Omega) & \text{if} \quad \Omega \neq \Omega_t^* \end{cases} \tag{3}$$

Again, if there is no unique maximum-valued theory but a set of two or more theories that yield equal values, the individual is assumed to choose randomly with equal probabilities from the theories in this set.[3] With these assumptions made, we can state

Lemma 2. *Let every individual decide on the choice of a theory according to the following rule: If* $\mathbf{W} := \{\Omega^n \in \{\Omega^1, ..., \Omega^N\} : \arg\max v(\Omega)\}$ *is a singleton, then the only* $\Omega^n \in \mathbf{W}$ *will be chosen. If* $|\mathbf{W}| = S > 1$*, then any* $\Omega^n \in \mathbf{W}$ *will be chosen with equal probability* $1/S$*. With individuals being heterogeneous with respect to* α*, the theory market will effectively collapse to two competing theories, the majority theory and a preferred minority theory, after a non-conformist choice has been made.*

Proof. The first individual at $t = 1$ decides randomly, so that some theory will become Ω^*. As long as the subsequent choices are made by individuals who are sufficiently conformist to also pick Ω^*, all other theories $\Omega \neq \Omega^*$ have a following of identical size. When the first individual enters who is sufficiently non-conformist to decide herself (randomly) for one particular $\Omega^m \neq \Omega^*$, then it follows straightforwardly from (3) that, due to its greater following, Ω^m will be preferred to any other

[3]Note that the modeling of conformism here is somewhat akin to that of habituation in the model on the formation of conventions by Hodgson and Knudsen (2004). There, individuals have an exogenous disposition to follow their habits, but their actual degree of habitually following a certain routine depends on their decisions of the past. Here, the individual of degree of conformism follows from some (maybe genetic) disposition, but the decision of a non-conformist to follow the majority or not depends on the actual size of that majority.

$\Omega \neq \Omega^*$. Therefore, individuals deciding after the first non-conformist choice will effectively always have the choice set $\{\Omega^*, \Omega^m\}$. □

All individuals who make a non-conformist decision at later stages of the process will also choose Ω^m, while all individuals making a conformist decision will choose the majority theory Ω^*. The market shares of all other theories will tend towards zero with more and more individuals deciding between Ω^* and Ω^m. The reason is that non-conformism in our model is defined only as a weak aversion against the majority group, which may even be overcome if the majority group becomes so large that it, by its sheer size, signals that it is unlikely that the opposition theory is reasonable. The black swan example above was one illustration, another may be the current debate about climate change. The sheer relative size of the group of those who believe in man-made climate change vis-à-vis the group of climate skeptics might be interpreted even by relative non-conformist individuals as a signal for a superior trustworthiness of the hypothesis that climate change is man-made. Were they confronted with a smaller majority, these moderate non-conformists would decide to join the camp of climate-skeptics.

As soon as a the theory market is collapsed to $N = 2$, (3) can be written as $(3a)$,

$$v(x) = \begin{cases} (1-\alpha) \cdot w(\Omega^*) \\ \alpha \cdot w(\Omega^m), \end{cases} \tag{3a}$$

and the α for which an individual is just indifferent between conformism and non-conformism can be calculated by equating both cases of $(3a)$, which yields

$$\bar{\alpha} = \frac{w(\Omega^*)}{w(\Omega^*) + w(\Omega^m)} \quad \text{with} \quad \lim_{t \to \infty} \bar{\alpha} = w(\Omega^*). \tag{4}$$

The convergence in time of $\bar{\alpha}$ towards $w(\Omega^*)$ follows simply from the fact that, once they are determined, only the majority theory and the preferred minority theory are chosen, so that the added market shares of these theories tend towards one.

That does not mean that the other $N-2$ theories that have existed on the theory market at $t=0$ disappear altogether, but they are marginalised and cease to have a noticeable impact on public discourse.

Knowing this, we can pin down a simple decision rule which connects the individual propensity to be a conformist with the the relative frequency of the two remaining theories on the theory market.

Lemma 3. *Sufficiently far into the sorting process on the theory market, individuals with $\alpha < \bar{\alpha} \approx w(\Omega^*)$ will strictly prefer Ω^* and individuals with $\alpha > \bar{\alpha} \approx w(\Omega^*)$ will strictly prefer Ω^m.*

Proof. Lemma 3 follows straightforwardly from combining (4) with Lemma 2. □

For those individuals who are exactly indifferent with $\alpha = \bar{\alpha}$, we will make the assumption that these citizens decide in a conformist fashion. To finally write down the q-function of the Polya-process discussed here, suppose that values of α (i.e., degrees of conformism) are normally distributed over the population with mean $\mu = 0.5$ and an arbitrary standard deviation σ. Given the simple decision rule, we must find, for any feasible value of w_t^* the probability that the next individual drawn from the population has an $\alpha \leq w_t^*$. Thus, given our assumption regarding the distribution of values of α in the population, we can then state that as soon as Ω^* and Ω^m are selected from the N available theories, the q-functions for these two theories converge towards

$$q_t^*(w_t^*) = \int_{-\infty}^{w_t^*} \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\frac{-(w_t^* - 0.5)^2}{2\sigma^2}} \quad dw_t^* \tag{5}$$

$$q_t^m(w_t^*) = 1 - q_t^*(w_t^*). \tag{6}$$

This leads to a characteristic sigmoidal graph for the two q-functions. Given that there is a positive probability that an individual has an $\alpha < 0$ or an $\alpha > 1$, it follows that $q_t^*(0) > 0$, $q_t^*(1) < 1$, $q_t^m(0) > 0$ and finally $q_t^m(1) < 1$. The exact

numerical values depend on σ; a rise of σ, would reflect a growing number of extreme conformists and nonconformists in the population. Such a change in the composition of the population is not modeled in this paper, however: σ is assumed to be constant.

The relationship between the actual fraction w_t^* and the probability q_t^* of the next individual also choosing Ω^* is depicted graphically in *Figure 1.* There are two stable

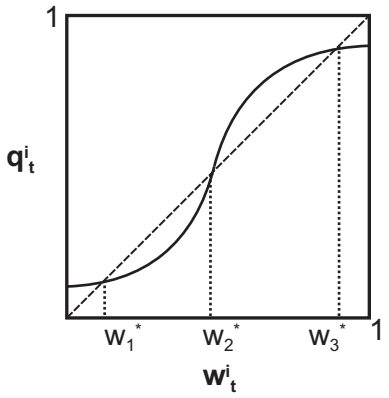

FIGURE 1. Equilibria on the theory market

equilibria for w^* on this theory market, one at w_1^* and one at w_3^*. In both cases, the probability of the next individual choosing Ω^* is higher than the actual fraction w_t^* for an interval around $w_{1,3}^*$ where $w_t^* < w_{1,3}^*$ and lower for an interval where $w_t^* > w_{1,3}^*$. The attracting intervals are delimited by the unstable equilibrium at $w_2^* = \mu = 0.5$. For any $w_t^* < w_2^*$, the process will converge towards w_1^* and for any $w_t^* > w_2^*$ it will converge towards w_3^*. Since Ω^* has been defined the majority theory at the outset, we can expect its market share to converge towards w_3^* without further interventions into the process; the market share of the preferred minority theory Ω^m will then converge towards $w^m = 1 - w_3^*$.

3. Interjurisdictional labour and capital markets

With the theory market being in place, we can now sketch a simple and very general two-jurisdiction economy, largely with standard properties. Note that the aim is deliberately *not* to use specific assumptions in order to guarantee unambiguous price

signals, but to preserve generality by imposing only a few restrictions on the model economy. Within this model economy, the market signals are produced that are ultimately processed on the political theory market.

3.1. **Loyal versus perfectly mobile individuals.** Suppose that every individual i can be characterised by additively seperable preferences for both income and policy,

$$U_i(x, y) = y_i + u(x) \tag{7}$$

where $u(x)$ is a positive procedural utility that can be gained from a policy vector x and that follows not from the outcome of a policy, but from the fact that the policy vector is in concurrence with the theory Ω^i held by the individual. The utility of the actual outcome of the imposed policy is captured by $y_i = l_i^T + k_i$, i.e. the sum of net wage and capital incomes, which will be determined below. Let $u(x)$ be a step function with $u(x) = \bar{u}$ if the individual lives in a jurisdiction where Ω^i determines policy and $u(x) = 0$ if not. Suppose further that an individual interested in other jurisdictions would need to invest at least amount of c to gather meaningful information on foreign income earning opportunities. However, if the jurisdictions are heterogeneous with regard to their majority theories, the utility $\bar{u}$ would be lost with a relocation. Assuming that i has an infinite time horizon and some a priori belief regarding the distribution of incomes that can be earned in other jurisdictions, at time $t = \tau$ she has an incentive to invest into gathering information about other jurisdictions only if

$$\sum_{t=\tau}^{\infty} \delta^{-(t-\tau)} \left[E(y^F) - y - u(x) \right] > c \tag{8}$$

where $\delta > 1$ is a discount factor, $E(y^F)$ is the expected value of income-earning opportunities to be discovered in foreign jurisdictions. It is evident from (8) that, with all other parameters fixed, there must exist a threshold level of political utility $\hat{u}$ where for all $\bar{u} \geq \hat{u}$ an individual abstains completely from gathering information

about foreign jurisdictions herself. On the other hand, we might also have $\bar{u} < \hat{u}$. These individuals will invest into gathering information about income earning opportunities in other jurisdictions and they will migrate whenever they find a jurisdiction where the income to be earned is sufficiently high. From these considerations, we arrive straightforwardly at

Lemma 4. *If individuals are sufficiently heterogeneous regarding either* $E(y^F) - y$*, or* c*, or* $u(x)$ *or all of these, the population can be separated into two subpopulations: One loyal group, whose members do not inform themselves about foreign jurisdictions at all, and one mobile group, whose members inform themselves and migrate whenever* $y^F - y$ *is sufficiently large.*

It is useful to further note that individuals who adhere to a minority theory can be expected to relocate whenever they find a foreign jurisdiction offering a $y^F > y$ because for them, $u(x) = 0$. Even more importantly, note also that not all loyal individuals need to be satisfied with the status quo policy in a jurisdiction. As long as the discounted expected income differential is smaller than the costs of searching, even politically disgruntled individuals will not actively pursue information on policy outcomes in other jurisdictions. But even if they do and indeed find a more lucrative jurisdiction, they will only actually relocate if the costs of migration are smaller than the discounted regional utility differential. Since these magnitudes need not be identical for all individuals, we can expect that some positive number of politically disgruntled individuals will remain in their home jurisdiction.

3.2. **A signal produced by decentralised policy.** To investigate the signals produced by factor migration, we introduce probably the simplest equilibrium conditions available in the literature on decentralised fiscal policy. We assume that each individual supplies one unit of homogeneous labor and is endowed with an arbitrarily high amount of homogeneous capital. Factors are allocated between two regions, A and B, with the private sectors in both regions being characterised by standard, neoclassical production functions. Adding to this, we assume that the vector $x = (\lambda, \theta, G)$ comprises the policy conducted by the public sector with θ

denoting a head tax, G denoting the quantity of a public good and λ denoting the technology used to provide the public good. Presuming a perfectly controlled government which frictionlessly enforces the majority's preferred policy in order to suppress control problems, the entire tax revenue is used to provide productive public goods and no rents are accrued by individuals in the public sector. Public policy enters the private sector production function through a function $\rho(x)$ with $\rho > 0 \forall x$. The effect of $\rho(x)$ is exactly the same as that of a Hicks-neutral, factor-augmenting public input, and accordingly the private sector production function is assumed to be linear-homogeneous. Thus, the complete production function for each of the two jurisdictions $j \in \{A, B\}$ is

$$Y = \rho(x_j) \cdot F(L_j, K_j). \tag{9}$$

Individuals are assumed to be uncertain regarding the function $\rho(x)$, and uncertainty here implies not only parameter uncertainty, but also uncertainty regarding the functional form of ρ – in other words, individuals act under model uncertainty and are compelled to act upon fallible hypotheses about the effects of policy changes on the aggregate output and on the marginal productivities of labour and capital. While the individuals know that $\frac{\partial \rho}{\partial G} > 0$ and $\frac{\partial \rho}{\partial \theta} < 0$, they do not know the exact functional form and can therefore not simply determine the optimal size of the public sector in a marginal calculus. Similarly, they know that a higher λ generates higher incomes than a lower λ, but gathering information about superior technologies of supplying public goods is, as already mentioned, costly.

Since we assume a perfectly controlled government and exclude rent-seeking activities, it is evident that the entire tax revenue is used to provide the public good G. The effective level of G, however, is assumed to also depend upon the technology of public good provision, which is represented by the technology parameter $\lambda > 0$, so that

$$G = \lambda \theta L. \tag{10}$$

In essence, the choice of policy can then be reduced to a choice of a tax rate θ and of a technology λ, with the level of public goods being fully determined by these parameters. In our context, the term "technology" is supposed to encompass a wide range of real-world phenomena: not only physical means of production, but also the composition of a portfolio of different types of public goods. For instance, a relatively low value for λ could signify an excessive emphasis on redistibutive activities compared to efficiency-enhancing public capital, whereas a relatively high value for λ signifies the opposite. This rather imprecise account of possible influences on λ mirrors the problem of model uncertainty that the individuals in the model face.[4]

Looking at regions that are effectively small open economies, let the return per unit of capital in every region be identical with the world interest rate, $k_j = \frac{\partial F}{\partial K} \cdot \rho(x) = r^*$. Furthermore, let labour be paid according to its marginal product, such that we have for the pre- and after-tax labour incomes

$$l_j = \rho(x_j)\frac{\partial F}{\partial L_j} \tag{11}$$

$$l_j^T = l_j - \theta_j. \tag{12}$$

It is important to note that a return per unit of capital at r^* does not imply that capital owners are fully informed about the effects of policy in other jurisdictions. One reason may be that, when deciding where to invest their capital, they observe only overall rates of return in other jurisdictions and cannot disentangle *ex ante* how large the effect of $\rho(x)$ on the rate of return in the foreign jurisdiction is. Another reason is that even if the effect of $\rho(x)$ were isolated, one would still need to invest substantial additional efforts into finding out the details of public policy – remember

[4]Since the policy-space is not one-dimensional here, involving the choice of λ and θ, this would traditionally contradict the stability of a median voter equilibrium and therefore the assumption of a tightly controlled government frictionlessly following the median preferences. Note, however, that stability here effectively comes from the theory market, where majority preferences are clearly defined in a stable equilibrium, i.e. λ and θ are determined unambiguously by the choice of an Ω.

that the simple technology parameter λ represents a possibly very complex real-world technology of public good provision.

Suppose for simplicity, and without loss of generality, that there are two jurisdictions. B is the relatively efficient region, i.e., the same amount of public goods is financed in B with a lower tax rate, or a higher amount of public goods is, due to a more advanced technology, financed with the same tax rate in B, compared to A. At the initial stage, before labour mobility is introduced, we thus have higher net incomes from labour in the low-tax jurisdiction B, $l_B^T > l_A^T$. This is the first type of signal produced by decentralised economic policy, namely a price signal, which sends two messages: (i) given the current policies, a positive number of units of labour could be utilised more efficiently in B than in A, and (ii) the policies in A and B lead to different incomes from supplying labour. This type of signal will henceforth be called a *differential signal*.

3.3. **A signal produced by factor mobility.** Recognition of wage differentials will induce migration of mobile (but not of loyal) individuals between A and B, and this will obviously have an effect on l_A^T and l_B^T. An equilibrium on the interregional labour and capital markets is reached when both conditions $l_A^T = l_B^T$ and $k_A = k_B = r^*$ hold simultaneously. However, wage differentials may be persistent in our model since, given *Lemma 4*, the share of mobile individuals is limited and migration might cease before wages are equalised. Also, note again that the sign of the impact of migration on l_A^T and l_B^T is deliberately not determined unambiguously in this model, as it is probably not fully determined in reality. This can, as we will see shortly, lead to Scenarios where migration of labour actually increases interjurisdictional wage differentials. Generally, differentiating (12) yields

$$\frac{\partial l_j^T}{\partial L_j} = \frac{\partial \rho(x_j)}{\partial L_j} \cdot l_j + \rho(x_j) \cdot \left[\frac{\partial l_j}{\partial L_j} + \frac{\partial l_j}{\partial K_j} \cdot \frac{\partial K_j}{\partial L_j} \right] \tag{13}$$

and there may exist intervals for L_j where the positive first term overcompensates a bracketed term that is negative on aggregate.[5] A migration of productive factors from the relatively inefficient region A to the relatively efficient region B then leads to a rise of l_A via the direct effect of out-migration on marginal productivity, but it also leads to less capital being used in A and to a decline of $\rho(x)$ via a loss of tax revenue. Similarly, an inflow of additional units of labour to B will lead to a decline of marginal labour productivity, but this may be overcompensated by positive effects generated by additional productive public input financed with an enlarged tax base. It is therefore not *ex ante* clear whether the net effect of migration on the net incomes in A and B will be positive or negative. This leads us to

Lemma 5. *The net effect of labour migration on the after-tax wages paid in both regions is generally ambiguous due to the interaction of a direct marginal productivity effect with an indirect tax base effect.*

Four different cases, which are summarised in *Table 1*, can result from migration between two jurisdictions.

Scenario	**I**	**II**	**III**	**IV**
$\frac{\partial l_A}{\partial L_A}$	>0	>0	<0	<0
$\frac{\partial l_B}{\partial L_B}$	>0	<0	>0	<0

Table 1: Possible reactions of wages to factor migration.

While the differential signal results from given prices, we observe here changing prices of labour in A and B resulting from a regional shifting of resources. Such signals associated with price changes will thus be called *shift signals*.

[5]Note that it is *not* feasible to use the capital market equilibrium condition, substitute $\rho(x) = \frac{r^*}{\partial F/\partial K}$ into (11) and determine that wages unambiguously decline with the amount of labour used in the production process. The reason is that with migration, $\rho(x)$ itself changes, at least due to tax base effects.

Note that a reliable equilibriating tendency is asscociated only with scenario IV. In scenario I, there is a clear disequilibriating tendency resulting from factor migration, and in the other two scenarios the existence of an equilibrium depends on the relative velocity of the income effects of migration. If the marginal effect of migration on marginal productivity and on the tax base in B are consistently smaller than that in A, then there will be a tendency towards an equilibrium, associated with higher (III) or lower (II) incomes in both A and B. However, given the fact that there is a group of immobile individuals in our model, factor migration can come to a rest even with persisting income differentials.

4. Learning from another jurisdiction's policies

4.1. **Collective learning on the theory market.** Picking up the thread of *Section 2.2*, some statements regarding the stability of an equilibrium on the theory market can be made. Let stability $s = s(\mu, t)$ be defined as the absolute number of individuals who simultaneously need to change their minds in order to transform the status quo majority theory Ω^* into a minority theory. Remember that, by definition, the equilibrium frequency of the majority theory Ω^* must be w_3^* – as long as it is indeed favored by the majority. If the value of μ, i.e. of the mean of the distribution of the non-conformism-parameter α, rises, the population becomes on average more non-conformist. As a result, the distance $w_3^* - w_2^*$ in *Figure 1* is reduced, so that the attracting region of w_3^* becomes smaller while the attracting region of w_1^*, where Ω^* is transformed into an equilibrium minority theory, is enlarged. On the other hand, the positive impact of time on the stability of an equilibrium follows from the technical assumption that at any point in time t exactly one individual decides which theory she wants to pick. As a result, the frequencies are stable in equilibrium, but the absolute number s of individuals that need to change their minds to move the theory market to another equilibrium rises with t. This may appear to be an unrealistic feature of the model on first sight, but it can also be argued that is helps to approximate the fact that long-conveyed, traditional modes

of thinking about economic policy are more difficult to change than theories that have only recently been introduced and that are not as deeply rooted.

Given the approach we have chosen to model the presence of a propensity to be conformist in the population, it is relatively straightforward to introduce a catalyst for collective learning processes. The basic question is: In which form can experience with status quo policies have an impact on the existence and the properties of particular fixed points. Remember that at the outset, before experience was considered, $\mu = 0.5$ was assumed, i.e., we assumed individuals to be symmetrically distributed along the lines of conformism and nonconformism. But it appears to be a plausible assumption that μ changes when, based upon experience, individuals start to have reason to believe that Ω^* is faulty.

If a policy based upon the majority theory produces disappointing results, we should expect that for individuals who still have to decide themselves, the propensity to be a conformist is reduced. The more implausible a theory appears in the light of evidence, the more difficult it becomes *ceteris paribus* to convince oneself that it is nevertheless the correct theory. The result is a decline of conformism in the following sense: An individual who is yet to decide on the theory to pick is characterised by a given value of α, and would thus have so far needed a certain relative frequency of $w(\alpha)$ associated with the majority theory to also become convinced of Ω^*. The more conformist the individual is, the smaller the necessary majority becomes, i.e. $\alpha \to 0$ implies $w(\alpha) \to 0.5$. But with rising skepticism and a decline of conformism in the yet-to-decide population, the same individual would experience an increase of his value of α – he would now only follow a larger, in this sense more convincing, majority. If such a large majority does not exist, he will instead become a believer in the minority theory. This leads us to

Proposition 1. *If there is a sufficiently large decline of mean conformism in the population (a sufficiently large increase of μ) for a sufficiently long transitory period of time (for sufficiently many individuals to choose a theory in this period of time) the status quo majority theory becomes the minority theory, and vice versa.*

Proof. Let $z = (w^*, q(w^*))$ be the status quo fixed point, associated with a frequency $w^* > 0.5$ for Ω^* and, given that z is a fixed point, $q(w^*) = w^*$. Let μ increase by an arbitrarily small, positive ε. This implies that $q_{\mu+\varepsilon}(w^*) < w^*$, and given our assumptions regarding the shape of $q(\cdot)$, z must wander to the left, $w^* - w^*_{\mu+\varepsilon} > 0$. Let $d = (w^d, q(w^d))$ denote the unstable fixed point which delimits the attracting regions of the majority fixed point of Ω^* from the attracting region of the second fixed point, where Ω^* would be transformed into the equilibrium minority theory. It is straightforward that $w^d = q(w^d) = \mu$ at all times, so that with any increase of μ, d shifts to the right, $w^d - w^d_{\mu+\varepsilon} < 0$. The result is that unequivocally, $\frac{\partial s(\mu,t)}{\partial \mu} < 0 \forall s > 0$. If we repeat this procedure often enough, we will at some stage approach $w^* - w^d = 0$. From thereon, and for any further increases of μ, the stable majority fixed point w^* ceases to exist. As long as μ does not fall below the critical value again, the relative frequency of Ω^* will deteriorate until a lower stable fixed point is reached, where Ω^* is now held only by a stable minority. If, after a period of transitory skepticism, μ bounces back to $\mu = 0.5$, the newly determined status of Ω^* as a minority theory is not in jeopardy, since the existence and stability of a minority fixed point follows directly from our assumption of a normally distributed α. □

The key to change on the theory market thus is, as one might have expected, disappointing experience with the policies that have been introduced as being in line with the status quo majority theory. It is, however, important to note that a substantial increase of non-conformism is necessary for this to occur. For only moderate increases, we will experience a decrease of the majority theory's market share, and an accompanying reduction of $s = s(\mu, t)$, but it will nevertheless remain the theory accepted by a majority of voters. We can think of the theories as being supported by very stable social networks in their fixed points, and these networks can absorb even relatively large shocks to which the theory market may be exposed. The remaining question thus is if, and under which condition, we can expect a shock that is sufficiently large to destabilise a given majority theory.

4.2. **Learning from differential signals.** What seems particularly appealing about learning from differential signals is that individuals can learn from the policies conducted in neighbouring jurisdictions without the occurrence of any potentially distorting spatial factor movements. In our framework, however, the prospects for political learning from differential signals are dim. We have seen in the first Proposition that a destabilising a given equilibrium on the theory market requires a substantial, widespread discontent with the status quo in the social communication processes about policy. Only in this case a sufficiently strong, transitory increase of non-conformism will occur to lead enough yet undecided individuals towards the preferred minority theory such that a reversal of equilibria on the theory market can take place. However, *Lemma 4* informs us that there is a group of loyal individuals in each jurisdiction who will not invest into gathering information on policy and political results in the other jurisdiction. Furthermore, these loyal individuals will be predominantly those who enjoy a high procedural-political utility u – those who support the incumbent majority theory. In other words, societal communication processes are dominated by those who support the status quo and who have an incentive to ignore, rather than spread information that casts doubt on the status quo.

Even the presence of a very strong and non-mobile non-conformist faction in the polity does not substantially change this argument. There may be such loyal non-conformists, if in (8) even with $u = 0$ the cost component assumes a larger value than the expected income differential. However, with the non-conformists being non-mobile, and acting behind a veil of insignificance in their domestic political process, they will not actively invest into receiving and (probably more importantly) correctly interpreting differential signals. Mobile non-conformists, on the other hand, migrate whenever they find a large enough income differential. They produce shift signals, but do not engage in the domestic discourse on policy. This reasoning leads us straightforwardly to

Proposition 2. *Differential signals will generally not suffice to destabilise an equilibrium on the theory market populated by non-mobile individuals acting behind a veil of insignificance.*

There is some empirical evidence, most prominently provided by Besley and Case (1995), on political yardstick competition. They show that voters in American states tend to deny re-election to incumbent governors who raise taxes while their colleagues in neighbouring jurisdictions do not, while they tend to accept tax raises when neighbouring governors also raise taxes. Voters appear to interpret a solitary tax increase as a sign of government waste, or rent-extraction, while a common tax increase is interpreted as a signal for a real economic necessity of increasing taxes. This observation supports our initial assumption of tightly controlled governments, but also seems to indicate that voters do in fact use the differential signal of tax rate differentials. However, the mechanism is shown to fail for other issues of economic policy: for regionally differing income levels and unemployment rates, Besley and Case find no significant influence on the individuals' voting decisions. This is a somewhat puzzling result; if voters learn from the comparison of regionally differing tax policies, and a lone tax raise is deciphered as a sign for inefficiency, then why does the same mechanism not work for other fields of policy? From the perspective of our model, two points can be made: (i) the tax rate alone is a very simple signal, and its isolated interpretation requires little effort compared to the interpretation of information on the interaction of λ, θ and G; (ii) voters appear to use this signal for controlling self-interested incumbents, but not to update their knowledge on economic policy as a whole.[6]

[6]Also, evidence on tax mimicking in other countries than the United States indicates that incumbent politicians anticipate the specific control function of yardstick competition with tax rates. As for instance Revelli (2001, 2002) shows for the United Kingdom, Heyndels and Vuchelen (1997) show for Belgium and Feld and Reulier (2005) show for Switzerland, tax increases in neighboring jurisdictions are often taken as an opportunity to increase taxes at home. For evidence on the relationship between decentralisation and control of representatives, see also Feld, Schaltegger and Schnellenbach (2007).

4.3. **Learning from shift signals.** The argument is substantially different for shift signals. Migration flows have an impact on the net incomes of the loyal individuals who remain in their home jurisdictions, and it is difficult to ignore a signal that manifests itself in a change of one's own budget constraint. Therefore, sufficiently large migration between jurisdictions, with sufficiently large decreases of the incomes of loyal individuals, can indeed be expected to lead to change the mood of societal communication processes on economic policy, i.e. to lead to an increase of skepticism and the mean of non-conformism in the population. However, Table 1 shows that, if one does not enforce restrictive assumptions, the sign of the effect of migration out of the relatively inefficient region on net wages is not fully determined. A larger tax base allowing to finance more productive infrastructure may overcompensate the direct effect on marginal labour productivity, or it may not. In the relatively inefficient region, net incomes may actually rise as a result of out-migration if the public goods effect does not overcompensate the direct effect on marginal productivity: Loyal individuals would then experience the shift signal as a positive loosening of their private budget constraint.

In Table 1, only Scenario I provides unambiguously efficient incentives. Migration out of the relatively inefficient and into the relatively efficient region leads to higher net wages in the relatively efficient, and lower net wages in the relatively inefficient region. As a result, in the region with the relatively inefficient policy, a distortion of the local theory market can eventually be achieved, and political experimentation can be conducted in this polity. However, from a standard neoclassical perspective, Scenario I is the most unlikely one, because in both jurisdictions the tax base effects would have to overcompensate the standard effects on wages in the production function. The other Scenarios are associated with perverse incentives for collective learning in at least one of the two jurisdictions. If Scenarios III or IV occur, the remaining individuals in the relatively inefficient region experience a raise of their incomes after labour mobility is implemented. In Scenario IV, the loyal individuals in the relatively efficient jurisdiction may become convinced to revise their relative efficient policy due to the "wrong" informational content of the price signal. We therefore have

Proposition 3. *Shift signals can be expected to destabilise equilibria on the theory markets if they are sufficiently strong (if migration flows are sufficiently broad), but there is no reason to systematically expect these signals to have the desired effect in, and only in, the relatively inefficient jurisdiction, where a change of policy would be unambiguously preferable.*

These problems could be reduced if individuals learn from both type of signals considered here. To also reckon that the level of incomes is higher in B than it is in A would certainly an improvement compared to an exclusive reliance on shift signals. This is especially true for the scenarios that imply a further divergence of income levels, i.e. that do not lead the model economy towards the equilibrium conditions on the interjurisdictional labour and capital markets, and where migration flows only end when all mobile individuals have exploited the regional income differentials. The additional information would enforce justified scepticism in A. If there is no divergence, though, then a convergence of income levels could easily serve as an argument to defend Ω_A^*: if the income level in B is decreasing, then this can be easily interpreted as an indicator for an increasing fitness of Ω_A^* relative to Ω_B^*. Thus, there is generally even more information necessary to ensure that individuals have the correct incentives. They would need to know the differential signal, the shift signal *and* they would need to reckon that the shift signal follows from migration and that migration out of A is a sign of relative inefficiency of Ω_A^*. This may be trivial for an economist – but for an individual who defends his set of conjectures behind a veil of insignificance and within a stabilising social network, a willingness to face the facts cannot be simply presupposed.

5. Conclusion and outlook

With all these arguments made, we can nevertheless conclude that from a very long-run, knowledge-producing perspective, decentralisation still ought to be preferred to centralisation. Decentralisation delivers a systematic tendency to destabilise equilibria on the theory market. Even if this does not necessarily occur in the correct (the inefficient) jurisdiction, a change of Ω_B^* would also produce new knowledge

about the efficacy of economic policies. A unitary system is missing the inherent instability that comes with the signals discussed here and is thus bound to produce less knowledge in the long run. However, we have shown that the problem exists that incentives to revise a given set of conjectures and thus to experiment with new policies are not necessarily to be found in the relatively inefficient region. While more knowledge is produced in decentralised systems, it cannot be ensured that there is a frictionless diffusion process where the relatively efficient policy is adopted by all jurisdictions.

Somewhat surprisingly, this result has also an encouraging facet, because diversity of policies is likely to be sustained. We are unlikely to experience *ex post* harmonisation towards one efficient policy, but rather an ongoing process where distorted equilibria on the theory market lead to a continuing revision of theories, which in turn leads to experiments with new policies. This is encouraging, because efficiency in the model is always only relative efficiency vis-à-vis other jurisdictions. In an ongoing process of distorted equilibria on the theory market, and experimentation with novel policies, there is always the prospect of discovering novel policies that are even more efficient.

Another conclusion is more disturbing from a normative point of view, although not surprising from an economic point of view. The famous, and highly controversial, Schumpeter hypothesis states that private entrepreneurs become more innovative, the larger their market share (see e.g. Witt 2002). A similar implication could be suspected here, since the tight control of representatives by rationally ignorant voters is a crucial element of our model. However, while private entrepreneurs, even if they have market power, are still disciplined to some degree through the price system, uncontrolled public entrepreneurs in pursuit of self-interested utility maximisation have a potential to make decisions that are immensely harmful to voters. Recommending to increase the leeway of representatives in order to facilitate political innovations would therefore be very premature and, in fact, dangerous (see Schnellenbach 2007 for a detailed discussion of this issue).

6. References

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1983): "A Generalized Urn Process and Its Applications", *Cybernetics* 19: 61-71.

Arthur, W. Brian, Yuri M. Ermoliev and Yuri M. Kaniovski (1987): "Path-Dependent Processes and the Emergence of Macro Structure", *European Journal of Operational Research* 30: 294-303.

Benz, Matthias and Alois Stutzer (2004): "Are Voters Better Informed when They Have a Larger Say in Politics?", *Public Choice* 119: 31-59.

Besley, Timothy and Anne Case (1995): "Incumbent Behaviour: Vote-Seeking, Tax-Setting and Yardstick Competition", *American Economic Review* 85: 25-45.

Feld, Lars P. and Emmanuelle Reulier (2005): *Strategic Tax Competition in Switzerland: Evidence from a Panel of the Swiss Cantons*, CESifo Working Paper 1516, Munich: CESifo.

Feld, Lars P. , Christoph Schaltegger and Jan Schnellenbach (2007): "On Government Centralization and Fiscal Referendums", forthcoming, *European Economic Review.*

Hayek, Friedrich A. von (1968): *Der Wettbewerb als Entdeckungsverfahren*, Kieler Vorträge, N.F. 56, Kiel: Institut für Weltwirtschaft.

Heyndels, Bruno and Jef Vuchelen (1997): "Tax Mimicking Among Belgian Municipalities", *National Tax Journal* 51: 89-101.

Hirschman, Albert O. (1989): "Having Opinions: One of the Elements of Well-Being?", *American Economic Review (Papers and Proceedings)* 79: 75-79.

Hodgson, Geoffrey M. (1997): "The Ubiquity of Habits and Rules", *Cambridge Journal of Economics* 21: 663-684.

Hodgson, Geoffrey M. and Thorbjørn Knudsen (2004): "The Complex Evolution of a Simple Traffic Convention. The Functions and Implications of Habit", *Journal of Economic Behavior and Organization* 54: 19-47.

Hodgson, Geoffrey M. and Thorbjørn Knudsen (2006): "Why We Need a Generalized Darwinism, and Why a Generalized Darwinism Is Not Enough", *Journal of Economic Behavior and Organization* 61: 1-19.

Kerber, Wolfgang and Nicole J. Saam (2001): "Competition as a Test of Hyptheses: Simulation of Knowledge-Generating Market Processes", *Journal of Artificial Societies and Social Simulation* 4, Issue 3.

Keynes, John M. (1936): *The General Theory of Employment, Interest and Money*, London: Macmillan.

Kollman, Ken, John H. Miller and Scott E. Page (2000): "Decentralization and the Search for Policy Solutions", *Journal of Law, Economics and Organization* 16: 102-108.

Kuran, Timur (1987): "Preference Falsification, Policy Continuity and Collective Conservativism", *Economic Journal* 97: 642-665.

Oates, Wallace E. (1999): "An Essay on Fiscal Federalism", *Journal of Economic Literature* 37: 1120-1149.

Revelli, Federico (2001): "Spatial Patterns in Local Taxation: Tax Mimicking or Error Mimicking?", *Applied Economics* 33: 1101-7.

Revelli, Federico (2002): "Testing the Tax Mimicking versus Expenditure Spill-over Hypothesis Using English Data", *Applied Economics* 34: 1723-32.

Rose-Ackerman, Susan (1980): "Risk Taking and Reelection: Does Federalism Promote Innovation?" *Journal of Legal Studies* 9: 593-616.

Salmon, Pierre (1987): "Decentralisation as an Incentive Scheme", *Oxford Review of Economic Policy* 3: 24-43.

Schnellenbach, Jan (2007): "Public Entrepreneurship and the Economics of Reform", *Journal of Institutional Economics* 3: 183-202.

Strumpf, Koleman S. (2002): "Does Government Decentralisation Increase Policy Innovation?", *Journal of Public Economic Theory* 4: 207-241.

Vanberg, Viktor A. and Wolfgang Kerber (1994): "Institutional Competition Among Jurisdictions: An Evolutionary Approach", *Constitutional Political Economy* 5: 193-219.

Vihanto, Martti (1992): "Competition Between Local Governments as a Discovery Procedure", *Journal of Institutional and Theoretical Economics* 148: 411-436.

Witt, Ulrich (1998): "Imagination and Leadership – The Neglected Dimension of an Evolutionary Theory of the Firm", *Journal of Economic Behavior and Organization* 35: 161-177.

Witt, Ulrich (2002): "How Evolutionary Is Schumpeter's Theory of Economic Development?", *Industry and Innovation* 9: 7-22.