Transaction Costs and Institutions*

Charles Nolan† Alex Trew‡

January 28, 2011

Abstract

This paper proposes a simple framework for understanding endogenous transaction costs – their composition, size and implications. In a model of diversification against risk, we distinguish between investments in institutions that facilitate exchange and the costs of conducting exchange itself. Institutional quality and market size are determined by the decisions of risk averse agents and conditions are discussed under which the efficient allocation may be decentralized. We highlight a number of differences with models where transaction costs are exogenous, including the implications for taxation and measurement issues.

JEL Classification: D02; D51; L14.
Keywords: Exchange costs, transaction costs, general equilibrium, institutions.

*We thank colleagues at the Universities of Edinburgh, Exeter, Koç and St. Andrews, and also at the RES Conference at Warwick University and the ISNIE Conference in Toronto. In particular, we thank, without implicating, Sumru Altug, Vladislav Damjanovic, John Driffill, Oliver Kirchkamp, John Moore, Helmut Rainer, József Sákovics, David Ulph and especially Tatiana Damjanovic, Daniel Danau and Gary Shea.

†Department of Economics, University of Glasgow. Tel: +44(0) 141 330 8693. Email: charles.nolan@glasgow.ac.uk.

‡Corresponding author: School of Economics and Finance, University of St. Andrews: Castlecliffe, The Scores, St. Andrews, Fife KY16 9AL, Scotland, UK. Tel: +44(0) 1334 461950. Email: alex.trew@st-andrews.ac.uk.

1 Introduction

Transaction costs are the costs of collecting information, bargaining, communicating, decision making, and enforcing contracts between individuals, firms and the state (Coase, 1960). The perception of such costs as exogenously impeding trade or inhibiting the formation of complete contracts suggests that reducing, eliminating or avoiding those exogenous costs is generally welfare enhancing.[1] As the quality of contracting institutions (institutions, for short) is thought to be a part of what explains those transaction costs,[2] the implication is that better institutions improve economic outcomes.

This paper introduces a simple framework in which the costs of exchange are determined by optimal investment in the quality of institutions. Total transaction costs are the sum of two components: There is a cost to forming the public and private institutions that govern, ex ante, the terms of exchange, and there are costs to *conducting* exchange once the state of nature is resolved.[3] Agents choose the resources allocated to reducing exchange costs *and* how extensively they will trade with others.

By viewing transaction costs as an endogenous component of a general equilibrium set-up, we are able to draw a number of important distinctions with models where transaction costs are left as exogenous. First, while the costs of exchange can be can be too high, they can also be *too low*. A high

[1]See, for example, Greenwood and Jovanovic (1990) and Townsend and Ueda (2006) in relation to finance, growth and inequality; Levchenko (2007) on international trade; and Dixit (1996) on political economy.

[2]See Levchenko (2007) and Acemoglu, Antràs and Helpman (2007) for a similar perspective.

[3]Throughout, an ‘exchange cost’ is the cost of conducting a particular exchange and what we refer to as the ‘transaction cost’ is the sum of investments in institutions and the subsequent costs of exchanges that occur.

exchange cost reflects fewer resources directed towards facilitating transactions but may be associated with greater expected utility if those resources increase consumption and agents choose to make fewer costly exchanges. Although this is a natural consequence of a set-up where we make transaction costs endogenous, we show that this has important implications for tax policy. Institutions might be designed to maximize the size of markets, for example, but when the costs of the institutions themselves are considered, we show that the simple market-size maximizing institution generally implies a higher than optimal tax rate. Moreover, an institution that minimizes transaction costs, such as suggested by the Transaction Cost Economics (TCE) literature,[4] is also seen to be generally sub-optimal.

Second, making transaction costs endogenous permits us to understand the existence and extent of what Wallis and North (1986) refer to as the transaction sector. Coase (1992, p.716) argues that "a large part" of economic activity is directed at alleviating transaction costs, whilst Wallis and North (1986) estimate that the transaction sector comprised half of US GNP in 1970, a proportion which had grown significantly over the preceding century. We find that a more wealthy economy is characterized by a smaller transaction sector, but one based on greater investments in institutions, larger markets and lower exchange costs.

Third, part of the contracting institution is naturally modelled as a pure public good and so for the competitive equilibrium to be efficient all political rents need to be competed away. Endogenous transaction costs thus suggest

[4]Equilibrium in the TCE framework is where "transactions... are aligned with governance structures... so as to effect a (mainly) transaction cost economizing outcome" (Williamson, 2010: p.681).

a rationale for the arguments of Schumpeter (1942) and Wittman (1989), that democratic *political* systems can deliver efficient *economic* outcomes.

Fourth, we consider the effect of a transactions tax. Where transaction costs are exogenous, agents respond by reducing market size for sufficient increases in tax. When they are endogenous, however, agents invest more in institutions to ameliorate the effect of the tax on the costs of exchange, thereby making diversification decisions less sensitive to increases in the transactions tax. However, while apparently more robust to the imposition of a transaction tax, agents opt for autarky at a lower transaction tax than might be anticipated.

Our framework relates to a number of other papers. The idea that investments in better transaction technologies can be part of an efficient economic system has been put forward by De Alessi (1983), Barzel (1985) and Williamson (1998). In making transaction costs endogenous, we are also blurring the distinction between institutional and technological efficiency, as Antràs and Rossi-Hansberg (2008) have noted in relation to organizations and trade. The gains from allocating resources to institutions relates to the idea of a 'state capacity' in Besley and Persson (Forthcoming); our approach also considers the possibility of a 'private capacity' that might substitute for or complement investments in public institutions. While this paper focuses on the costs incurred in risk sharing, Martins-da-Rocha and Vailakis (2010) considers how transaction costs in financial markets can be made endogenous to the labor costs incurred in financial intermediation.

The model is set out in Section 2. Section 3 characterizes efficient equilibria, that is the optimal investment in institutional capital, the optimal extent of transaction costs and the optimal market size. The conditions are discussed in Section 4 under which the efficient equilibrium may be decentralized. Section 5 reports the implications of our model in the light of extant empirical evidence. Section 6 examines the impact of (simple) non-optimal institutions while Section 7 looks at the impact of an exogenous transaction tax. Section 8 summarizes and concludes. Appendices contain some proofs, further numerical analysis of the model and a detailed description of the decentralized equilibrium.

2 The model: overview

We briefly outline the model (which is motivated by Townsend, 1978) before presenting it in detail. A large number of risk averse agents are each endowed with the same amount of capital. Agents can differ in the technology with which they can produce the consumption good but, initially, know only the distribution of possible technologies. Consumption risk can be reduced by forming markets with other agents but due to exchange costs it is not feasible to replicate a complete Arrow-Debreu allocation. The cost of each bilateral exchange is determined in a simple way by the quality of contracting institutions in the economy.[5] As such, *before* they realise their productivity, agents decide whether or not to form a market with other agents and, if they do, how large that market should be and how much to invest in private (i.e., excludable) and public institutional quality.

[5] Intuitively, consider the adoption, ex ante, of standard accounting practices by a set of firms. It is doubtless costly to establish such a framework. However, ex post, there are still costs to running the system such as inputting data or prudential auditing.

One agent per market becomes an intermediary, buying outputs and selling consumption bundles. Intermediaries are here the productive unit of the transaction sector, using the institutional capital as input to a common 'exchange cost technology' (ECT), the output of which is the exchange cost incurred by agents in its market. Ex post, agents honor their obligations even if it would be preferable to renege. If agents do not join a market then no institutional investment takes place. The primary aim of Sections 3-4 is to characterize the efficient level of institutional capital, exchange costs and market size (consumption risk-sharing) for such a model economy.

2.1 Preferences and production technology

The economy is populated by a countable infinity of agents, $i \in I$. All agents have the same utility function, $u(c)$, with a constant degree of relative risk aversion. Each is endowed with the same amount of capital, $0 < k < \infty$. Agent i produces amount $\lambda^i y^i$ of the non-storable consumption good, where $y^i \leq k$ is the amount of capital used in production. The set of possible technologies, Λ, is finite and bounded away from infinity. λ^i is distributed i.i.d. across agents with $p\left(\lambda^i\right)$ denoting the probability of any agent drawing λ^i, and $\sum_{\lambda\in\Lambda} p(\lambda) = 1$.

Let ω represent the state of nature, i.e., a list of λ^i for all $i \in I$. Let Ω be the set of all possible states of nature, and $p(\omega)$ the probability of some ω occurring and $\int_{\omega\in\Omega} p(\omega) = 1$.[6]

[6] Precisely, $p(\omega)$ is the probability of the state of nature in a small interval $d\omega$ occurring.

2.2 Diversification and intermediation

To diversify against risk, agents can form markets in a star network around a single intermediary, as in Townsend (1978). A *market* is a set of agents $M \subset I$ with cardinality $\#M < \aleph_0$; so a market is finite-sized. The set of agents with whom agent i exchanges directly is denoted N^i. Agent $h \in M$ is an *intermediary* if $N^i = h$ for $i \in M \backslash h$ and $N^h = M \backslash h$. So markets are disjoint and agents only exchange with one intermediary. M^h denotes the market intermediated by agent $h \in H$ where $H \subseteq I$ is the set of all intermediaries in the economy.

Agents can exchange some of their endowment of capital on a one-for-one basis for shares in the consumption portfolio compiled by the intermediary, less their contributions to the total costs of transactions. Each agent in a market exchanges twice with the intermediary, so for a given exchange cost, α, the per-agent cost of exchange in a market of size $\#M$ is $2\alpha\left(\frac{\#M-1}{\#M}\right)$.

2.3 Institutional capital and exchange costs

Intermediaries form markets using institutional capital as an input to an exchange cost technology (ECT) which determines the cost of exchange in that market. There is free-entry to intermediation (i.e., the ECT is accessible to all agents). Institutional capital takes two forms: There is a market-specific institutional capital (S-capital) and a general economy-wide institutional capital (G-capital). Agent i allocates the portion τ_s^i and τ_g^i of the endowment to each respective type of capital, where $0 \leq \tau_s^i + \tau_g^i \leq 1$. In a market intermediated by agent h, the exchange cost, α^h, is determined as follows:

$$\alpha^h = \left[1 - F\left(S^h, G\right)\right] k. \tag{1}$$

where $F()$ is the ECT, which is concave, continuous and increasing in each of its arguments and satisfies some Inada-type conditions.[7] The range of F is the unit interval, so $0 \leq \alpha^h \leq k$. Since the exchange cost is proportional to k it is akin to an 'iceberg cost'. We assume that S-capital is the average contribution of agents in a market and G-capital is the average across the whole economy.[8] Given that all agents are identical ex ante, all will make the same allocations and so we generally omit superscripts. Clearly, then, $S \leq \tau_s k$ and $G \leq \tau_g k$. The ex ante allocations, $(\tau_s + \tau_g) k$, combined with the ex post costs, $2\alpha\left(\frac{\#M-1}{\#M}\right)$, make up total transaction costs.

The distinction between S-capital and G-capital is natural. S-capital reflects activities specific to the transaction being made–organizational choices, private infrastructure or learning about property rights–and is local and excludable. On the other hand, G-capital reflects the general legal enforcement of contracts, fiduciary duties, public infrastructures or competition policy but is, by contrast, an economy-wide public good. Each type of institutional capital clearly impacts upon the other, and we assume that S and G are complements, so $F_{SG} = F_{GS} > 0$. Without a public institution, S-capital would be less effective; without S-capital, there may be little point in an

[7] Namely: $F(0,0) = 0$ and $F(S,G) \to 1$ as $\{S,G\} \to \{\infty,\infty\}$. Next, $F_S(S,G) \to \infty$ as $S \to 0$, and $F_S(S,G) \to 0$ as $S \to \infty$; analogous conditions obtain with respect to G. So defined, F ensures that the 'null' arrangement of zero investment in S-capital and G-capital is always equivalent to there being no gains from trade.

[8] In other words, there is no *institutional* scale effect: The introduction of an agent into a market requires an equal additional contribution to keep the exchange cost for that market constant.

economy-wide institution.

3 Efficient equilibria

The purpose of this Section is to characterize efficient allocations in a cooperative setting; i.e., in an economy where intermediaries arise exogenously and where there is no difficulty providing the public good, G. We subsequently show that the efficient cooperative allocations coincide with those that are in the core (Proposition 5), and then that competitive equilibria are Pareto optimal under some conditions on the provision of G (Section 4).

An intermediary h makes an allocation for a market which is an n-tuple $\{\#M, \{\tau_s, \tau_g\}, \alpha, y, c^i(\omega)\}$. The intermediary allocation is that which maximises his own expected utility,

$$\max_{\#M,\tau_s,\tau_g,\alpha} \left\{ \int_{\omega\in\Omega} p(\omega)\, u\left[c^h(\omega)\right] \right\} \tag{2}$$

subject to,

$$\begin{aligned}
\sum_{i\in M} c^i(\omega) &\leq \sum_{i\in M} \lambda^i(\omega) y && (3)\\
2\alpha(\#M-1) &\leq \#M\left[(1-\tau_s-\tau_g)k - y\right]; && (4)\\
G &\leq \tau_g k; && (5)\\
S &\leq \tau_s k; && (6)\\
\tau_s+\tau_g &\leq 1; \tau_s \geq 0; \tau_g \geq 0; && (7)\\
\alpha &= \left[1-F(S,G)\right]k; && (8)\\
E\left[u\left(c^i\right)|h\right] &\geq E\left[u\left(c^i|h'\right)\right] \quad \forall i, \forall h',\ \text{for } h' \text{ feasible.} && (9)
\end{aligned}$$

Equation (3) limits consumption to be less than or equal to output in each state of nature, on a market-wide basis. Equation (4) says that, in any market, the sum of endowments net of production and institutional investment must be sufficient to cover the total of market exchange costs. Equations (5)-(6) describe institutional capital, (7) restricts the range of feasible $\{\tau_s, \tau_g\}$, and (8) describes the exchange cost. Finally, equation (9) is a participation constraint requiring that the utility of all agents in market h be at least as high as participating in another feasible market.[9] Implicit in the maximisation of (2) is a participation constraint for the intermediary; if,

$$\sum_{\lambda\in\Lambda} p(\lambda)\, u\left[\lambda k\right] > \int_{\omega\in\Omega} p(\omega)\, u\left[c^h(\omega)|h\right], \text{ for all } h \text{ feasible,}$$

then no intermediated market with exchange is formed (i.e., $\#M = 1$, $\tau_s = \tau_g = 0$).

3.1 Optimality

All agents share an aliquot consumption payout in each market as determined by the observed average technology for that market, $\bar{\lambda}(\omega, h) = (\#M)^{-1}\sum_{i\in M^h} \lambda^i(\omega)$. One more assumption is required before stating Proposition 1 which establishes that well-defined solutions to (2) exist. Recall that total per capita costs are $2\alpha\left(\frac{\#M-1}{\#M}\right)$. If for all k it is the case that $(1-\tau_s-\tau_g)\,k \leq 2\alpha$, then the optimal market size will be bounded since perfect diversification would imply non-positive consumption. Hence, let an arbitrary maximum value of k be $\bar{k}$. Let $K \equiv [\underline{k}, \bar{k}]$ where $\underline{k} < \bar{k}$, be the set of potential endow-

[9] It will turn out that (7) is implied by (9).

ments. When $k = \bar{k}$, as shown below, investments in ECT will be at their maximum level: $\bar{\tau}_s, \bar{\tau}_g$.[10] A sufficient condition for finite-sized markets to be part of the optimal plan is then,

$$2(1 - \max F(\cdot)) \geq (1 - \bar{\tau}_s - \bar{\tau}_g). \tag{10}$$

(10) ensures that there exists a $\bar{M} \geq \#M(\bar{k})$, where $\bar{M}$ is a finite integer which may serve as an upper bound on optimal market size. Let $\mathbb{M} \equiv \left[1, ..., \bar{M}\right]$ be the set of feasible market sizes. In short, (10) imposes that exchange costs do not fall 'too quickly' as institutional capital rises.

Proposition 1 *For each* (k, ω) , *(i) the maximum of (2) is attained and (ii) the value function,* $V(k)$, *is well-defined and continuous. (iii) The optimal policy correspondence may not be unique.*

Proof. The relevant measure space is given by the triple (Ω, ω, p), where ω is a σ-algebra, the collection of all the subsets of Ω, and p is the measure defined on ω. An agent's expected utility is, therefore, $Eu(c) \equiv \int_{\omega \in \Omega} p(\omega) u\left(c\left(\omega\right)\right)$, where the utility function is strictly concave. Also, note that:

$$k \in K \subset R_{++}; \tag{11}$$

$$\{\tau_s, \tau_g\} \in T \subset R_{+}^{2}; \tag{12}$$

$$\#M \in \mathbb{M} \subset \mathbb{N}. \tag{13}$$

[10] S-capital and G-capital are essentially normal goods, 'purchases' of both rising in k.

Decisions on τ_s, τ_g and $\#M$ are taken after k is known but before agents' productivities are revealed. Using (12) and (13), define the feasible policy choices as follows: $\Gamma : K \rightarrow T \times \mathbb{M}$. That is, $T \times \mathbb{M}$ is the product space,

$$\left\{ \left\langle [0,1]^2, 1 \right\rangle, \left\langle [0,1]^2, 2 \right\rangle, \left\langle [0,1]^2, 3 \right\rangle, ..., \left\langle [0,1]^2, \bar{M} \right\rangle \right\}.$$

For each $\omega \in \Omega$, Γ is clearly a non-empty, compact-valued, continuous correspondence. A typical element of that mapping is denoted $z \in \Gamma(k, \omega)$. The optimization problem, therefore, involves a strictly concave criterion function and a non-empty, compact constraint set, so that the maximum is attained. Since the maximum is attained, the value function, $V(k)$, is well-defined. It follows from the Theorem of the Maximum that $V(k)$ is continuous. As the feasible policy set is not strictly convex, the optimal policy

$$G(k, \omega) = \{z \in \Gamma(k, \omega) : Eu(c) = V(k)\}$$

need not be unique. Finally, the equivalence between the cooperative case and core allocations is established in Proposition 5. ■

Some properties of efficient equilibria are immediate. Since the ECT is freely accessible, rents from intermediation must be zero so (5)–(6) hold with equality. The optimal allocations to the ECT is the pair $\left\{ \tau_s^*, \tau_g^* \right\}$ which satisfies,

$$\theta k F_S = 1; \tag{14}$$

$$\theta k F_G = 1, \tag{15}$$

where $\theta \equiv 2\left(\frac{\#M-1}{\#M}\right)$, and F_X $(X \equiv S, G)$ denotes a partial derivative of $F\left(S^h, G\right)$. So, for any market size, the optimal institutional investments equate the marginal cost with the marginal gain from reducing the exchange cost. When there are no transactions, agents would not undertake such investment and the optimal choice of market size satisfies

$$\arg\max_{\theta} Eu((1-\theta)\,k\lambda). \tag{16}$$

Clearly, the optimal market size is unity, $\theta = 0$. The choices $\tau_s = 0$, $\tau_g = 0$ and $\#M = 1$ also define the unique reservation utility for any agent to be a member of any $\#M > 1$ coalition, $\overline{V} = \sum_{\lambda \in \Lambda} p(\lambda)\, u[\lambda k]$, $\forall k$. That furthermore ensure that $0 \leq (\tau_s + \tau_g) < 1$.[11] A direct implication of the reservation value of utility is that any equilibrium of the model in which $\tau_s > 0$ and $\tau_g > 0$ is one in which $\#M > 1$, necessarily.

Remark 1 follows from analysis of (14) and (15):

Remark 1 *For $\#M > 1$ and endowment k, Equations (14) and (15) define a unique pair, $\left\{\tau_s^{\#M}, \tau_g^{\#M}\right\}$ such that $\tau_s^{\#M} > 0$ and $\tau_g^{\#M} > 0$. For $\#M = 1$, equations (14) and (15) imply $\tau_s^{\#M} = 0$ and $\tau_g^{\#M} = 0$. For a given endowment, optimal investments in the ECT rise and bilateral exchange costs fall as the market size increases.*

[11] That (16) generates a unique reservation value of utility is not quite so trivial as it may at first appear. In particular, it is different to Townsend (1978). In that paper, agents take the per capita exchange cost as given. Even so, the optimal market size may well be greater than one. In the present set-up, when agents take the exchange cost as given, the optimal market size is necessarily one. This is due to our scaling the exchange cost by capital.

The choice of any market size greater than one is determined by the ECT and agents' attitude to risk. (14) and (15) determine efficient investment in the ECT and hence total resources diverted from goods production. Agents' risk aversion provides an upper bound on how much they are willing to pay for consumption insurance, given an alternative not to diversify; that bound is independent of the ECT. Agents will optimally form markets with $\#M > 1$ when *efficient* investment in the ECT delivers transaction costs lower than that bound.[12]

The rest of the analysis of equilibrium is contained in the following four Propositions. One can show that in general there is a critical level of k below which transaction costs dominate and agents do not diversify:

Proposition 2 *There is a $k > 0$ small enough such that optimal market size is one.*

Proof. See Appendix A. ■

The key to understanding that Proposition is to recall that $S = \tau_s k$. So, a given τ has a larger proportionate impact on α as k rises, even though α itself is proportional to k. Hence, higher levels of k permit lower ex post transaction costs and help sustain larger market sizes and the benefits from consumption risk sharing.

Proposition 2 also indicates the possibility of multiple optimal plans. Proposition 3 now shows that no more than two such alternative plans can exist.

[12] That explains why agents may move from 'autarky' ($\#M = 1$) to a market size $\#M > 2$ for a small change in k. This is apparent in the numerical simulations – diverisficiation at market sizes in between is too costly given the ECT and the degree of risk aversion. Proposition 3 shows that subsequent increases in market size will be in steps of one.

Proposition 3 *The optimal policy correspondence contains at most two plans.*

Proof. See Appendix A. ■

As a corollary, it follows that for any $\#M > 1$ which is optimal, subsequent market size increases are single steps for (sufficiently large) increases in k. The final Proposition and following remarks establish that multiple equilibria can arise although they are, in a sense, special cases.

Proposition 4 *Let $\mathcal{K}$ denote the Borel sets of $K \subset R_{++}$. There exists a level of $k \in \mathcal{K}$, call it k^*, such that*

$$Eu\left\{\left[\left(1-\tau_1\right)k^* - \theta_1\alpha_1 k^*\right]\bar{\lambda}\left(\omega, M_1^h\right)\right\} = Eu\left\{\left[\left(1-\tau_2\right)k^* - \theta_2\alpha_2 k^*\right]\bar{\lambda}\left(\omega, M_2^h\right)\right\}$$

where $\tau_1 \neq \tau_2, \alpha_1 \neq \alpha_2$, $M_1^h \neq M_2^h$ and where maximized utility is identical under both programs.

Proof. See Appendix A. ■

The set of values of k which result in multiple equilibria is of measure zero. That is, such values of k^* correspond to pairs of θ's; these $\theta's$ are a proper subset of the rationals and hence themselves drawn from a set of measure zero.

To summarize: Low endowment economies may resort to autarky (Proposition 2). However, for economies with larger endowments ($k > \underline{k}$) it is optimal to invest in the ECT and to form markets (Proposition 1 and Remark 1). Equilibrium plans need not be unique (Proposition 4) but those equilibria are, in a sense, of limited interest (Proposition 3 and the brief discussion following Proposition 4). Finally:

Proposition 5 *The allocations of the cooperative economy coincide with core allocations.*

Proof. See Appendix A. ■

4 Equilibrium with competitive intermediation

The question now is: Can the efficient outcome be decentralized? First, note that given the public good nature of the G-capital, there is a free-riding problem associated with investment in G-capital:

Proposition 6 *In the core, voluntary allocations to general capital are zero.*

Proof. See Appendix B. ■

In arguments later echoed forcefully by Wittman (1989), Schumpeter (1942, p. 269) asserted that political markets work in much the same way as do competitive economic markets: "(T)he democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote."[13] We formalize this perspective and consider a democratic political outcome to be the result of a political market with free-entry: Any agent can costlessly propose a manifesto which specifies taxes for all agents and the level of G-capital to be provided. The manifesto with the largest

[13] Wittman (1989, pp. 1395–6) argues "...that democratic political markets are organized to promote wealth-maximizing outcomes, that these markets are highly competitive and that political entrepreneurs are rewarded for efficient behavior."

share of votes is enforced ex post.[14] Proposition 7 shows that such *political* competition then yields an efficient outcome; the competitive G provision will be $G^* = \tau_g^* k$.

Proposition 7 *If any agent may propose an enforceable political manifesto then all agents will be taxed according to (15). It follows that $G = G^*$.*

Proof. See Appendix B. ■

There is no such problem in the provision of S-capital as free-entry to intermediation (i.e., any agent may become an intermediary) ensures that all rents are competed away. In Appendix C the decentralized economy is studied. It is established that there exists a unique, Pareto-optimal allocation in each market with a unique intermediary in each market.

Proposition 8 *The provision of S-capital and G-capital is Pareto optimal in the competitive equilibrium.*

Proof. See Appendix C. ■

5 The size of the transaction sector

Wallis and North (1986) were the first to quantify the size of the US transaction sector. They found that in 1870 it accounted for 25% of GNP, rising to 50% in 1970.[15] That the size of the transaction sector appears positively

[14] So we invoke ex ante perfect competition and ex post monopoly in the political process. Relaxation of either assumption might provide a focus for understanding the existence and power of elites.

[15] The transaction sector appears to be important in other advanced economies. See Wang (2003) and Klaes (2008) for surveys.

related to wealth has been thought to sit uneasily with the view that low transaction costs are one of the spurs of development.[16]

Our simple framework is not designed to connect directly with the empirical estimates of the transaction sector and a number of problems exist in an attempt to do so.[17] Nevertheless, the model developed here may be used to reinterpret the Wallis and North finding. We let κ measure the size of the transaction sector (the proportion of the endowment not allocated to the production of the consumption good):

$$\kappa \equiv (k - y)/k = \tau_s + \tau_g + \theta \left[1 - F(S, G)\right] \tag{17}$$

Some comparative static results are intuitive. Suppose that at some k, the optimal market size is $\#M' > 1$; κ is necessarily higher at any $\#M'' > \#M'$ since otherwise an equilibrium with lower transaction costs and better risk-sharing properties is available. The higher is k, the lower is κ for any given market size since agents invest more to reduce the costs of exchange (see Appendix D). As k increases, then, the cost of further diversification approaches the additional consumption-smoothing gain and for a sufficiently large increase in k optimal market size increases.[18] Figure 1 summarizes

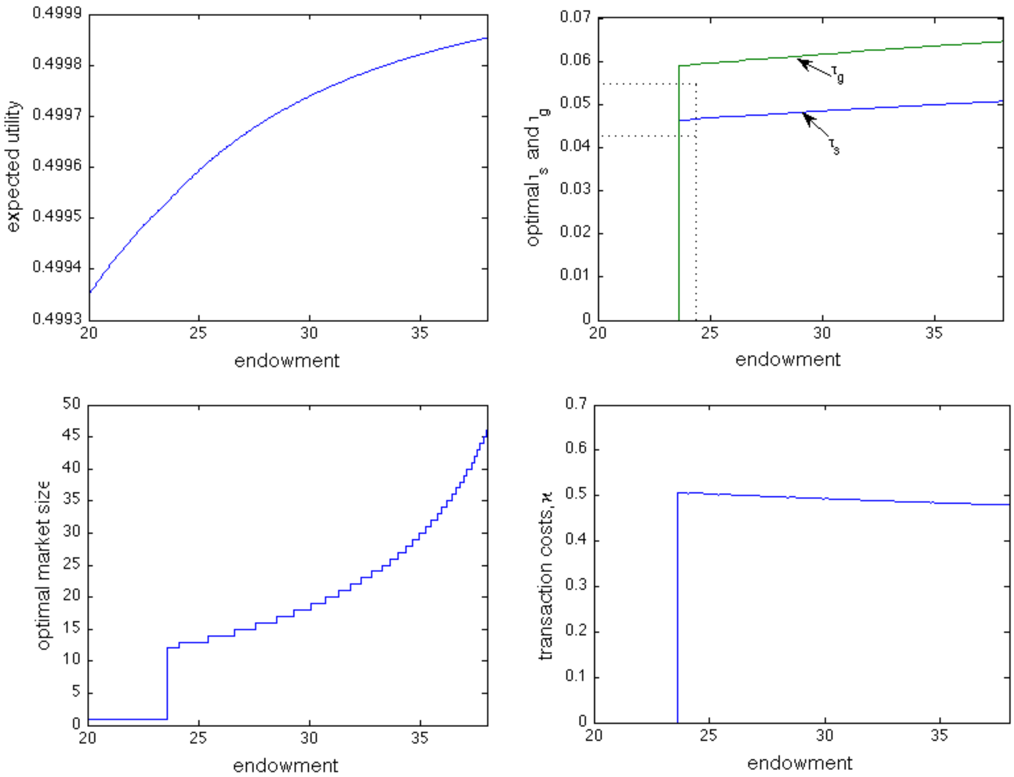
[16] Consider the conclusions in Klaes (2008): "...economies with less well-developed transaction sectors appear to exhibit lower levels of transaction costs if those costs are measured in terms of sector size, whereas micro-structurally those economies in fact suffer from higher levels of transaction costs due to significant barriers to smooth exchange and coordination of economic activity."

[17] The model hinges on a distinction between ex ante and ex post costs which, empirically, is not generally measured. Moreover, we assume agents are acting optimally.

[18] Proposition 3 shows that market size must increase by one from any $\#M > 1$, and the discussion preceding Proposition 2 explains how market size can jump by more than one when moving from $\#M = 1$.

results from a numerical version of the model (see Appendix E).[19] One finds that optimal transaction costs can be substantial; that large part of market activity referred to by Coase (1992) is reflected in our simple general equilibrium framework.

Figure 1: Equilibrium over a range of the endowment



Proportionate investments in institutions can increase in k, as Figure 1 demonstrates;[20] while the number of exchanges increases and institutional

[19] We adopt an ECT with constant elasticity of substitution of 2 whilst agents have a coefficient of relative risk aversion of 3. For Figure 1 we allow the effectiveness of S-capital and G-capital to differ slightly, although none of our results are sensitive to this assumption. See Appendix Table 2 for details.

[20] Appendix D identifies a sufficient condition on the ECT for that to be the case.

investment grows, total transaction costs ultimately fall (although the relation is non-monotonic). In other words, a wealthier economy has a larger and more *sophisticated* transaction sector, one based on greater investments in contracting institutions, larger markets and lower exchange costs. Empirically, the *payoff* from the greater investments – lower costs of individual exchanges – are difficult to identify in sectoral analyses which then mistake the higher investments in institutions as higher transaction costs overall. What appears to be obscured in the Wallis and North analysis is the distinction between ex ante investments in institutional capital and the ex post cost of conducting exchange; empirical analyses which make that distinction would appear worthwhile.

6 Distortive transaction cost policies

So far the analysis has focused on the nature of efficient equilibria; the efficient tax level, τ_g^*, provides an optimal trade-off between the expected level of consumption and its variability. This Section considers the impact of distortive (i.e., non-optimal) levels of the tax, τ_g.[21] First, we look at constrained optimal decisions about τ_s and $\#M$ over a range of imposed tax rates. Second, we consider the welfare implications of the policies implied by models of exogenous transaction costs.

[21] In this section S-capital and G-capital are equally efficient in reducing transaction costs so any difference between τ_s and τ_g reflects the institutional distortion. Specifically, $\gamma_s = \gamma_g = 0.075$ and $k = 30$. All other parameters are as in Appendix Table 2 unless otherwise stated.

6.1 Distortive institutions

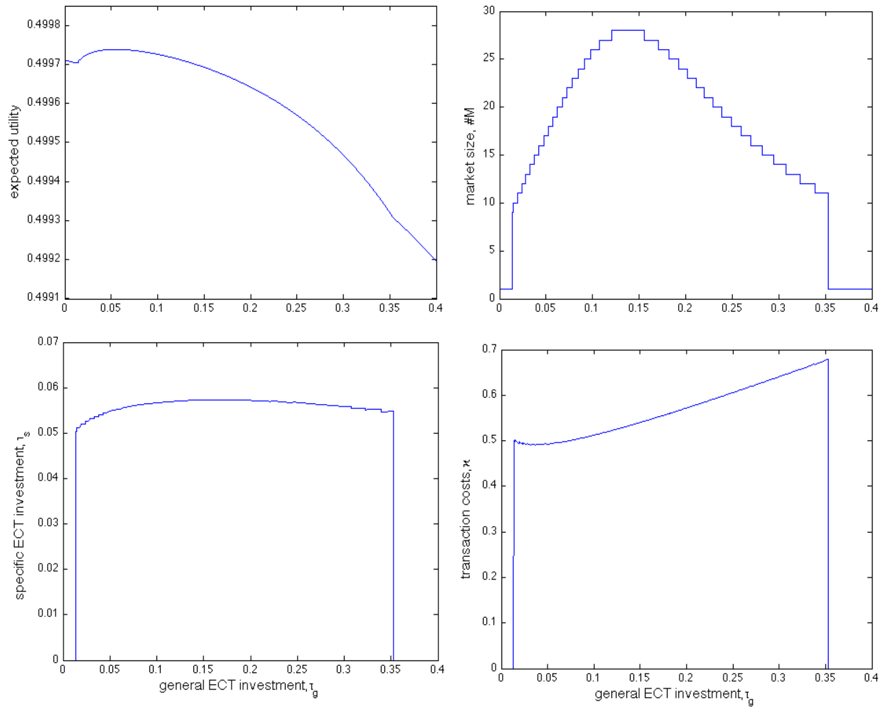
The distortive behavior centres on the resources turned over to G-capital. Such distortions might reflect deeper political tensions, perhaps between the electorate and a 'political class'. Alternatively, they may reflect 'irrational' voters, as in Caplan (2008). In any event, we do not model explicitly the origins of these tensions.

Deviations in τ_g from its efficient level can result in significant compensating changes in agent behavior, as Figure 2 shows. When τ_g deviates from τ_g^* agents can respond by changing τ_s and/or market size. If market size does not change, then τ_s always varies positively with τ_g. This follows from (14) and the assumption that G and S are complements; a higher G increases the marginal return to S-capital investment. Market size may rise or fall, for a sufficient increase in τ_g. A higher τ_g lowers the exchange cost and so, initially, makes further risk-sharing (increasing market size) attractive enough to incur the cost of a higher κ. However, for $\tau_g \gg \tau_g^*$, diverting more resources from production is costly and the gain from further diversification is low; as such, market size and τ_s both fall because that is the only way to reduce κ. In short, market size is hump-shaped in τ_g and, for some low and high values of τ_g, agents optimally choose not to diversify (setting $\tau_s = 0$ and $\#M = 1$), even though their endowments are being taxed.

6.2 Transaction cost policies

There are at least three types of institutional objectives suggested by models with exogenous exchange costs: First, institutions should deliver zero ex-

Figure 2: Equilibrium with distortive institutions



change costs; second, institutions should maximise risk-sharing/market size; and, third, institutions should minimize the size of the transaction sector. In our model, zero exchange costs are not feasible; institutions which deliver an infinite number of trades at zero cost are themselves infinitely costly. However, the analysis above shows that we can consider the second and third type of policy prescription. The second is a rule that maximizes the (constrained optimal) choice of market size; the third minimizes the size of the transaction sector. The market size policy is given by,

$$\tau_g^M := \{\min \{\tau_g | \#M \geq \#M'\} \text{ and } \tau_s \text{ optimal}\} . \tag{18}$$

That is, τ_g^M is the lowest tax required to induce the maximum market size, given that agents optimally choose market size and τ_s in response to τ_g. The tax policy which minimizes the transaction sector is given by,

$$\tau_g^\kappa := \min\{\kappa | \#M > 1 \text{ and } \tau_s \text{ optimal}\}. \tag{19}$$

For each policy the percentage change in certainty equivalent consumption is calculated, e.g., $\left[1 - u^{-1}\left(u^M\right)/u^{-1}\left(u^*\right)\right] \times 100$, in moving to τ_g^M from τ_g^*. Table 1 describes various features of equilibrium under the different rules.

Table 1: The Effects of Distortive Institutions.

rule	$\#M$	κ	τ_g	τ_s	α	$\%C$
τ_g^*	18	0.4940	0.0550	0.0550	6.0795	-
τ_g^M	28	0.5234	0.1210	0.0571	5.3711	4.3368
τ_g^κ	14	0.4906	0.0376	0.0537	6.4502	0.6400

Neither rule is optimal in a framework with endogenous transaction costs; the τ_g^M rule delivers too little *directly* productive capital and τ_g^κ too much consumption variability. It is useful to analyze which is the more costly, and why. Consider first the τ_g^M rule. Larger markets means, in short, higher taxation to lower the costs of bilateral exchange. The transaction sector is larger as a whole and its *composition* has shifted toward ex ante investments. τ_g cannot increase by too much, however, since agents always have the option of shifting resources to goods production by reducing market size and τ_s. For the τ_g^κ rule, minimizing transaction costs requires exchange costs to be too *high*. A low τ_g means that agents optimally form smaller markets and make smaller investment in S-capital, both actions reducing κ.

The loss from τ_g^M is greater than that from τ_g^κ.[22] Institutions fostering smaller exchange costs and bigger markets appear to be the more damaging in welfare terms. By targeting market size, the government 'distorts' choices of both $\#M$ *and* τ_g, leaving agents, in effect, with only one instrument to respond, τ_s. The alternative τ_g^κ, which minimizes the sum of all resources allocated to transactions, leaves private agents with the choice of market size and τ_s. The option for agents simply to consume their own endowment constrains policy such that the outcome is not too far from that which maximizes expected utility. That additional flexibility appears to reflect some of the empirical findings of Acemoglu and Johnson (2005) who argue that agents may find a variety of formal and informal mechanisms to obviate poor contracting institutions.

7 Transaction taxes

Policies are sometimes designed to raise revenue from a sector characterized by high-frequency trading or to reduce the number of trades where those trades are thought to be socially harmful in some way. We can consider a tax, t, that simply makes individual exchanges more expensive,

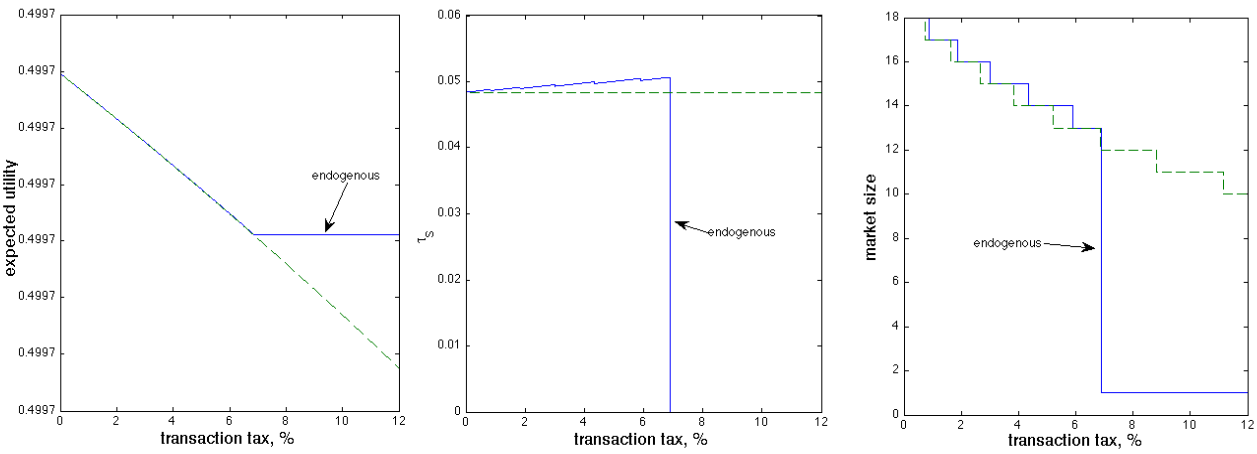
$$\alpha = (1+t)\left[1 - F(S,G)\right]k. \tag{20}$$

Figure 3 demonstrates the effect of a transaction tax where transaction

[22]This welfare ranking is robust to different values of k.

costs are endogenous (solid line). Agents respond first by increasing investments in institutions in order to dampen the effect of the tax on the cost of exchange. Relative to an environment where transaction costs are exogenous (dashed line),[23] market size is more robust to the introduction of a transaction tax. In each environment, agents can avoid the tax by simply resorting to autarky whenever the participation constraint is no longer satisfied. The important difference when transaction costs are endogenous is that agents have individually allocated a portion of their endowment to support institutions; in the exogenous case this investment has not occurred and so cannot be retrieved by agents. That means that although agents appear less affected by the tax, they will opt for autarky at a level of the tax far lower than that suggested when transaction costs are considered exogenous.[24]

Figure 3: The effects of a transaction tax



[23]The exogenous transaction cost set-up fixes α and the net endowment such that the equilibrium when $t = 0$ is the same as that in the endogenous transaction cost environment (i.e., τ_g and τ_s are fixed at the values which are optimal for $t = 0$).

[24]With endogenous transaction costs, autarky occurs at 6.9%; if we take them to be exogenous, the autarky equilibrium is induced at 22.2%.

8 Discussion and concluding remarks

Economists are increasingly focusing on the role of good institutions in promoting growth, trade and other desiderata. The intention of this paper is to link in a simple way impediments to transactions, institutional quality and market size. The efficient equilibrium of the model is consistent with significant transaction costs and investments in institutions. A distinction between transaction costs and exchange costs was made. The impact of distortive institutions was also considered, although in a tentative and ultimately ad hoc way. We argued that a number of what might be thought of as 'good' institutions are actually sub-optimal when transaction costs are endogenous.

It would seem important to extend the analysis in a number of directions. First, although agents had different productive capabilities, this had a limited impact as decisions over τ_s and τ_g were made *before* types were revealed. If decisions over τ_s and τ_g were made *after* agents' productive capabilities were known (and capabilities are private information) then the analysis will be somewhat more complicated. Related to this, incentive compatibility issues were not to the fore because it was assumed that agents remained in markets even if, ex post, they might have been better off under autarky. Nevertheless, the framework developed above may prove useful in the analysis of optimal tax and the role of government.

References

Acemoglu, D., Antràs, P. and Helpman, E. (2007). 'Contracts and Technology Adoption', *American Economic Review* **97**(3): 916–943.

Acemoglu, D. and Johnson, S. (2005). 'Unbundling Institutions', *Journal of Political Economy* **113**(5): 949–95.

Antràs, P. and Rossi-Hansberg, E. (2008). 'Organizations and Trade', NBER Working Paper No. 14262.

Barzel, Y. (1985). 'Transaction Costs: Are They Just Costs?', *Journal of Institutional and Theoretical Economics* **141**: 4–16.

Besley, T. and Persson, T. (Forthcoming). 'State Capacity, Conflict and Development', *Econometrica* .

Caplan, B. (2008). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*, Princeton University Press.

Coase, R. H. (1960). 'The Problem of Social Cost', *Journal of Law and Economics* **3**: 1–44.

Coase, R. H. (1992). 'The Institutional Structure of Production', *American Economic Review* **82**(4): 713–9.

De Alessi, L. (1983). 'Property Rights, Transaction Costs, and X-Efficiency: An Essay in Economic Theory', *American Economic Review* **73**(1): 64–81.

Dixit, A. (1996). *The Making of Economic Policy: A Transaction-Cost Perspective*, MIT Press.

Greenwood, J. and Jovanovic, B. (1990). 'Financial Development, Growth and the Distribution of Income', *Journal of Political Economy* **98**(5): 1076–1107.

Klaes, M. (2008). 'The History of Transaction Costs', *The New Palgrave Dictionary of Economics, 2nd Edition* . Durlauf, S. N. and Blume. L. E. (eds.), Palgrave Macmillan.

Levchenko, A. A. (2007). 'Institutional Quality and International Trade', *Review of Economic Studies* **74**(3): 791–819.

Martins-da-Rocha, V. and Vailakis, Y. (2010). 'Financial markets with endogenous transaction costs', *Economic Theory* **45**: 65–97.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy,* Harper & Row, New York.

Townsend, R. M. (1978). 'Intermediation with Costly Bilateral Exchange', *Review of Economic Studies* **55**(3): 417–25.

Townsend, R. M. and Ueda, K. (2006). 'Financial Deepening, Inequality, and Growth: A Model-Based Quantitative Evaluation', *Review of Economic Studies* **73**(1): 251–293.

Wallis, J. J. and North, D. C. (1986). 'Measuring the Transaction Sector in the American Economy, 1870–1970'. Chapter 3 in Engerman, S. L. and Gallman, R. E. (eds.), *Long-Term Factors in American Economic Growth* . University of Chicago Press.

Wang, N. (2003). 'Measuring Transaction Costs: An Incomplete Survey'. Ronald Coase Institute, Working Paper Number 2.

Williamson, O. E. (1998). 'Transaction Cost Economics: How it works; where it is headed', *De Economist* **146**(1): 23–58.

Williamson, O. E. (2010). 'Transaction Cost Economics: The Natural Progression', *American Economic Review* **100**: 673–90.

Wittman, D. (1989). 'Why Democracies Produce Efficient Results', *Journal of Political Economy* **97**(6): 1395–1424.

A Proofs of Section 3 propositions

Proposition 2 *There is a $k > 0$ small enough such that optimal market size is one.*
Proof. To the contrary, assume that

$$Eu\left(\left[\left(1-\tau_s-\tau_g\right)k-\theta\alpha\right]\bar{\lambda}\left(\omega, M^h\right)\right) \geq Eu\left(\lambda^i k\right), \tag{21}$$

for all k. All variables on the left-hand side reflect optimizing decisions. Note, in particular, that $\theta > 1, \forall k$. Observe that as $k \to 0$, $(1-\tau_s-\tau_g)-\theta(1-F(\cdot)) \to (1-\theta) < 0$, and that $\bar{V}$ is positive for all $k > 0$. Thus, there exists a level of k, call it $\underline{k}$, such that

$$\lim_{k\to\underline{k}} Eu\left(\left[\left(1-\tau_s-\tau_g\right)k-\theta\alpha\right]\bar{\lambda}\left(\omega, M^h\right)\right) \to 0,$$

As $k \to \underline{k}$, the inequality in (21) is reversed. ■

Proposition 3 *The optimal policy correspondence contains at most two plans.*
Proof. Let k^* denote a value of k such that $Eu(P_1|k^*) = Eu(P_3|k^*) = V(k^*)$. Assume that there also exists a P_2 such that $Eu(P_1|k^*) = Eu(P_2|k^*) = Eu(P_3|k^*)$. Let

$$\tau_3 > \tau_2 > \tau_1;\ \theta_3 > \theta_2 > \theta_1.$$

Now, denote

$$\theta_3 = \frac{x}{y}; \theta_2 = \frac{p}{q}; \theta_1 = \frac{m}{n}, \tag{22}$$

where x, y, p, q, m and n are all positive integers. Define χ as follows:

$$\chi = \frac{\frac{p}{q}-\frac{x}{y}}{\frac{m}{n}-\frac{x}{y}}.$$

By assumption $\frac{p}{q}-\frac{x}{y} < 0$, $\frac{m}{n}-\frac{x}{y} < 0$, and $\frac{p}{q}-\frac{m}{n} < 0$, so $1 \geq \chi \geq 0$. Finally, note that $q = p+1$, $n = m+1$ and $y = x+1$. Thus

$$\frac{p}{q} = \left[\frac{p(p+1)^{-1}-x(x+1)^{-1}}{m(m+1)^{-1}-x(x+1)^{-1}}\right]\frac{m}{n} + \left[\frac{m(m+1)^{-1}-p(p+1)^{-1}}{m(m+1)^{-1}-x(x+1)^{-1}}\right]\frac{x}{y}; \tag{23}$$

that is, market size associated with P_2 is a weighted average of the other two optimal market sizes. Hence it follows, by the strict concavity of the utility function, that either: *(i)* P_2 is indeed an optimal plan and P_1 and P_3 are not; or, *(ii)* P_2 is optimal and identical to either P_1 or P_3; or, *(iii)* $P_2 = P_1 = P_3$.

$\blacksquare$

Proposition 4 *Let K denote the Borel sets of $K \subset R_{++}$. There exists a level of $k \in K$, call it k^*, such that*

$$Eu\left\{\left[(1-\tau_1)\,k^* - \theta_1\alpha_1 k^*\right]\bar{\lambda}\left(\omega, M_1^h\right)\right\} = Eu\left\{\left[(1-\tau_2)\,k^* - \theta_2\alpha_2 k^*\right]\bar{\lambda}\left(\omega, M_2^h\right)\right\}$$

where $\tau_1 \neq \tau_2, \alpha_1 \neq \alpha_2$, $M_1^h \neq M_2^h$ and where maximized utility is identical under both programs.

Proof. Let $k \in [\underline{k}, \bar{k}]$. Partition that set into into $[\underline{k}, \hat{k})$ and $(\hat{k}, \bar{k}]$ such that the $\max Eu(\cdot \mid \forall k < \hat{k}) < \max Eu(\cdot \mid \forall k > \hat{k})$, and where $\tau(k > \hat{k}) > \tau(k < \hat{k})$, $\#M(k > \hat{k}) > \#M(k < \hat{k})$. By Proposition 2, such a partition is possible; it is also implied by (14)–(15). Let $\{k_1\}$ denote any sequence in $[\underline{k}, \hat{k})$ converging to $\hat{k}$ and let $\{k_2\}$ be any sequence in $(\hat{k}, \bar{k}]$ converging to $\hat{k}$. Let $V^1(k_1)$ denote $\sup Eu(\cdot|k \in [\underline{k}, \hat{k}))$, $V^2(k_2)$ denote $\sup Eu(\cdot|k \in (\hat{k}, \bar{k}])$ and $\hat{V}(\hat{k})$ denote $\sup Eu(\cdot|k_1 = \hat{k})$. By Proposition 1 these value functions are well defined and, by the Theorem of the Maximum, continuous. Hence, there exists a δ such that for $\left|k_1 - \hat{k}\right| < \delta/2$ and $\left|k_2 - \hat{k}\right| < \delta/2$ one has that,

$$\begin{aligned}
\left|V^1(k_1) - \hat{V}(\hat{k})\right| &< \varepsilon/2; \\
\left|V^2(k_2) - \hat{V}(\hat{k})\right| &< \varepsilon/2.
\end{aligned}$$

Hence

$$\begin{aligned}
\left|V^1(k_1) - V^2(k_2)\right| &\leq \left|V^1(k_1) - \hat{V}(\hat{k})\right| + \left|V^2(k_2) - \hat{V}(\hat{k})\right| \\
&\leq \varepsilon/2 + \varepsilon/2 \\
&= \varepsilon.
\end{aligned}$$

for k_1, k_2 close to $\hat{k}$ market size will not be changing. Hence, market size and taxes are higher for all $k \in (\hat{k}, \bar{k}]$ compared with $k \in [\underline{k}, \hat{k})$. Expected utility is identical with different optimal plans at $k^* = \hat{k}$. $\blacksquare$

Following Townsend (1978) and Boyd and Prescott (1985) one may characterize core allocations directly. In the discussion of the efficient equilibrium in the main text we studied the equilibrium decision rules of an intermediary. Townsend (1978) labelled that analysis the "cooperative" solution. Hence, the equivalence of the core and cooperative solutions is now established, Proposition 5 in the text.

Proposition 5 *The allocations of the cooperative economy coincide with core* *allocations.*

Proof. Consider the unique equilibrium: $x^h = \{c^*, y^*, \tau^*, \#M^*; i \in M^h\}$ for all $h \in H$. This n-tuple determines $F(S^*, G^*)$ and hence α^*. Suppose a strict subset of agents in the market intermediated by h, $B \subset M^h$, can form a blocking coalition (i.e., market). The cooperative equilibrium necessitates that the blocking coalition cannot deviate from contributing τ_g^* on average. The blocking intermediary chooses τ_s^B given the market size $\#B < \#M^h$. The agents in the blocking coalitions are better off with the following program:

$$\begin{aligned} c^i &= \bar{c}, \qquad \tau_g^i = \tau_g^*, \qquad \forall i \in B; \\ 1 &= 2\left(\frac{\#B - 1}{\#B}\right) kF_S\left(S^B, G^*\right). \end{aligned}$$

The consumption profile follows from optimizing over agents with identical, strictly concave utility functions and the second condition was derived in the text. By Remark 1, S^B-capital is strictly lower and exchange costs strictly higher. There are fewer transactions in this proposed market but each is more costly. In addition, the investment portfolio is less diversified. Given that market size and investment in S-capital deviate from the optimum, it must be that the higher exchange cost and less diversification are not compensated by fewer transactions; expected utility is necessarily lower. Now consider the case $B \supset M$. The same argument applies: In this proposed market, exchange costs are strictly lower and investments higher. The deviation from first-best means that expected utility must be lower than in market M. Hence, the cooperative allocation is in the core. Further, since the core is non-empty and $\#M$ is the unique optimal market size, it follows that the core allocations and the allocations of the cooperative economy coincide. ■

B Proofs of Section 4 propositions

Proposition 6 *In the core, voluntary allocations to general capital are zero.*

Proof. Consider an equilibrium in which $\tau_g^i = 0 \; \forall \; i \in I$, so that $G = 0$. By Proposition 2, for k big enough, there is a positive level of S-capital that is optimal, $\tau_s^i = \tau_s^{**}$, $\forall \; i \in I$; $EU^i = EU_0$ denotes expected utility under this plan. Suppose a blocking coalition B exists such that an agent $b \in B$ proposes $\tau_g^b > 0$ for $i \in B$. If $\#B < \aleph_0$, then $G = 0$ obtains and it follows that $EU^i < EU_0$ for all $i \in B$. Suppose, however that $\#B = \aleph_0$ and that an agent $b \in B$ proposes $\tau_g^b > 0$ for $i \in B$. In that case $G = G^b > 0$. Since $\#B > 0$, some positive level of G is optimal, and so $EU^i > EU_0$ for each

$i \in B$. But now there exists a blocking coalition $B' \subset B$ in which some agent $b' \in B'$ proposes $\tau_g^{b'} = 0$ for each $i \in B'$. Since $\#B' < \aleph_0$ it remains the case that $G = G^b$. Therefore, EU^i (for $i \in B'$) must be greater than EU^i (for $i \in B$). So while there can exist blocking coalitions which propose $\tau_g^b > 0$ for some $i \in B$, they are not in the core; 'voluntary' contributions to G-capital are zero. ■

Proposition 7 *If any agent may propose an enforceable political manifesto then all agents will be taxed according to (15). It follows that $G = G^*$.*

Proof. Consider that, in the process of electioneering, a group of agents $\mathcal{V} \subseteq I$ offer manifestos to be voted upon. Let $\mathcal{G} = \tau_g k$. The manifesto $\mathcal{M}^g = \{\{\tau^{ig}\}_{i\in I} | G \leq \mathcal{G}\}$ of each agent $g \in \mathcal{V}$ includes taxation levels for each agent $i \in I$ as well as a proposed level of G-capital, $G \leq \mathcal{G}$. It is assumed that there is no deception; agents deliver on their manifesto, if 'elected'. Agent i votes for the manifesto that will provide her with the highest expected utility; $EU^i|_{g'}$ is the expected utility to agent i if agent g' forms the government. $V^{g'}$ denotes the set of agents who vote for the manifesto of agent g'. So, if $\#V^{g'} > \#V^{g''}$ for every $g'' \in \mathcal{V} \backslash g'$, then agent g' forms the government, imposes taxation levels and delivers the level of G-capital in return. In the core, all agents are taxed equally and no rents accrue: $G = \mathcal{G}$. Consider the alternative to this. If an agent $g' \in \mathcal{V}$ offers a manifesto $\mathcal{M}^{g'} = \{\{\tau^{ig'} = \tau_g'\}_{i\in I} | \mathcal{G}'\}$ in which $\tau_g' k > \mathcal{G}'$ there is some other agent $g'' \in \mathcal{V}$ who offers a manifesto $\mathcal{M}^{g''} = \{\{\tau^{ig''} = \tau_g'\}_{i\in I} | \mathcal{G}''\}$ in which $\tau_g' k > \mathcal{G}'' > \mathcal{G}'$ which delivers $\#V^{g''} > \#V^{g'}$. Given 'perfect competition' in the political process, the rent from governing is driven to zero, so that agent g^*, who sets τ_g^* to satisfy (15), ensures that $G = \mathcal{G} = G^*$. ■

C Proof of proposition 8

The equivalence of core and competitive equilibria is established by extending the arguments of Townsend (1978). First, some notation is developed. Let any agent $h \in I$ propose strategy P^h for intermediating in a market. This strategy has eight components: M^h is the market proposed by agent h; P_1^h is the yield in terms of the consumption good of one share in the portfolio of agent h; P_2^h is the price in terms of the capital good at which agent h is willing to buy an unlimited number of shares in the goods production of any agent in M^h; P_3^h is a fixed fee in terms of the capital good for the purchase of shares in the portfolio of agent h by $i \in M^h$; P_4^h is the price in terms of the capital good at which agent h is willing to sell an unlimited number of shares in her portfolio to agents in M^h; τ_s^h is the proportion of the capital endowment that agent h proposes to invest in S-capital for that market; τ_g^h

is similarly defined. Recall that with free political entry agent g^* delivers on the manifesto promise and ensures that $G = \mathcal{G}$. Finally, $\alpha^h(F)$ is the ex post exchange cost. It is not strictly necessary to include $\alpha^h(F)$ in the definition of the strategy space but it aids intuition to do so. In what follows, Q_D^{ih} is the quantity of shares purchased by i in h's portfolio, whilst Q_S^{hi} is the quantity of shares sold by i to h. A^{ih} is a switching function, where $A^{ih} = 1$ if agent i buys shares in the portfolio of intermediary h, and $A^{ih} = 0$ otherwise. One may characterize the optimal strategies for intermediaries and non-intermediaries for a given market size. The following unconstrained optimization delivers the supporting price vector. Finally, we reduce on notation by writing x when we really mean $x(\omega)$.

Definition 1 *A competitive equilibrium is a set of actions* $\left\{Q_{S*}^{ij}, Q_{D*}^{ji}, A_*^{ij}\right\}$ *and a strategy* $P_*^i = \left\{M_*^i, P_{1*}^i, P_{2*}^i, P_{3*}^i, P_{4*}^i, \tau_{s*}^i, \tau_{g*}^i, \alpha_*^i(F)\right\}$ *for each agent* $i \in I$ *(where for any variable* x*, the convention* $x^{ii} = 0$ *is adopted), an allocation* $\{c_*^i, y_*^i; i \in I\}$ *and a set of markets which satisfy:*

1. If agent i is not an intermediary ($Q_{D*}^{ji} = 0, \forall j \in I$) then $\left\{y_*^i, Q_{S*}^{ij}, Q_{D*}^{ji}, A_*^{ih}, \tau_{s*}^{hi}, \tau_{g*}^{hi}, \alpha_*^{hi}(F)\right\}$ maximizes (24) subject to (25), (26) and (27) and $P^h = P_*^h \; \forall h \neq i$. c_*^i is given by (26), and given (30) – (33).

2. Agents participate in, at most, one market, M. In each market there is one intermediary such that $h \in M$, $M_*^h = M$, and $\left\{\tau_s^{ih}, \tau_g^{ih}, \alpha^{ih}(F)\right\} = \left\{\tau_{s*}^{ih}, \tau_{g*}^{ih}, \alpha_*^{ih}(F)\right\}$. For each $i \in M \backslash h$, $A^{ih} = 1$. For every such h, P^h is feasible, with y^h chosen to maximize (28) subject to (29) and (30). c_*^h is given by (30).

3. There exist no blocking strategies for any agent of I.

Hence, a non-intermediary faces the following problem:

$$\max_{Q_D^{ih}, Q_S^{hi}, y^i, A^{ih}} \sum_{h: i \in M^h} \left[A^{ih}\right] \left[EU\left(Q_D^{ih} P_1^h + y^i \lambda^i - Q_S^{hi} \lambda^i\right)\right], \tag{24}$$

subject to the following constraints regarding feasibility and participation,

$$\left[\sum_{h:i\in M^h}\left(A^{ih}\right)\left((1-\tau_s^h-\tau_g^h)k^i+Q_S^{hi}P_2^h-Q_D^{ih}P_4^h-P_3^h-\alpha(F^h)-y^i\right)\right]\geq 0; \tag{25}$$

$$c^i(\omega)=\left[\sum_{h:i\in M^h}\left(A^{ih}\right)\left(Q_D^{ih}P_1^h(\omega)+y^i\lambda^i(\omega)-Q_S^{hi}\lambda^i(\omega)\right)\right]\geq 0,\forall\omega\in\Omega; \tag{26}$$

$$A^{ih}=1\Rightarrow A^{ij}=0,\forall i \text{ such that } i\in M^j,\forall j\neq h. \tag{27}$$

An intermediary chooses a strategy to maximize,

$$\max EU\left(y^h\lambda^h+\sum_{i\in M^h}\left[Q_S^{hi}\lambda^i-Q_D^{ih}P_1^h\right]\right), \tag{28}$$

subject to,

$$\left[(1-\tau_s^h-\tau_g^h)k^h-\alpha(F^h)\left(\#M-1\right)-y^h+\sum_{i\in M^h}\left[Q_D^{ih}P_4^h+P_3^h-Q_S^{hi}P_2^h\right]\right]\geq 0; \tag{29}$$

$$c^h(\omega)=\left[\sum_{i\in M^h}\left[Q_S^{hi}\lambda^i(\omega)-Q_D^{ih}P_1^h(\omega)\right]+y^h\lambda^h(\omega)\right]\geq 0,\forall\omega\in\Omega; \tag{30}$$

and the following relations:

$$G \leq \tau_g k; \tag{31}$$

$$S^h \leq \tau_s k; \tag{32}$$

$$\alpha^h = \left[1-F\left(S^h,G\right)\right]k; \tag{33}$$

$$\sum_{i\in M^h}P_3^h-\alpha\left(\#M-1\right) = -2\alpha\frac{(\#M-1)}{\#M}. \tag{34}$$

Equation (34) may be regarded as the free-entry constraint upon the equilibrium intermediary strategy.

Proposition 9 *If there exists a competitive equilibrium, with political free-entry as defined above, then the equilibrium is in the core.*

Proof. Consider an equilibrium, for a given endowment k, for which a blocking coalition exists. Further, assume that this blocking coalition, denoted $\#B$, is a market in which transaction costs, production and consumption are Pareto optimal. In such an equilibrium, consumption is equalized across all agents in the blocking coalition: $y = (1 - \tau_s - \tau_g)\, k - \left(\frac{\#B^b - 1}{\#B^b}\right) 2\alpha^b$ and $c^b = \bar{c}, \forall i \in B$; $EU(c^b) > EU(c^h)$, $h \neq b$. For some agent $b \in B$, the following are necessary conditions for an optimal strategy:

$$\begin{aligned}
U'(\cdot) \sum_{i \in B^b} \lambda^i + \phi^b\, (\#M)\, P_2^b + \mu^b \sum_{i \in B^b} \lambda^i &= 0; && (35)\\
-U'(\cdot)\, (\#B)\, P_1^b - \phi^b\, (\#B)\, P_4^b - \mu^b\, (\#B)\, P_1^b &= 0; && (36)\\
U'(\cdot)\lambda^b + \phi^b + \mu^b \lambda^b &= 0; && (37)
\end{aligned}$$

$$\phi^b k - \phi^b F'(S^b, \mathcal{G})(\#B-1)k^2 - \eta^b F'(S^b, \mathcal{G})(\#B-1)k^2 + 2\eta^b F'(S^b, \mathcal{G})k^2 \frac{(\#B-1)}{\#B} = 0, \tag{38}$$

where ϕ^b, μ^b, η^b, are unknown multipliers on constraints (29), (30) and (34), respectively. For non-intermediaries, necessary conditions for an optimum are,

$$\begin{aligned}
U'(\cdot)\lambda^i + \phi + \mu\lambda^i &= 0; && (39)\\
-U'(\cdot)\lambda^i - P_2^b\phi - \mu\lambda^i &= 0; && (40)\\
U'(\cdot)P_1^b + \phi P_4^b + \mu P_1^b &= 0, && (41)
\end{aligned}$$

where ϕ and μ are multipliers on constraints (25) and (26), respectively. Equations (39) and (40) together imply that $P_2^b = 1$, and so it follows that $P_4^b = 1$. Use these in the intermediary's first-order conditions:

$$P_1^b = -\frac{\phi^B}{U'(\cdot) + \mu^B}, \tag{42}$$

and,

$$U'(\cdot) = -\frac{\phi^h\, (\#B)\, P_2^b}{\sum_{i \in B^b} \lambda^i} - \mu^b. \tag{43}$$

Therefore,

$$P_1^b = \left(\frac{1}{\#B}\right) \sum_{i \in B^b} \lambda^i. \tag{44}$$

Finally, we have,

$$P_3^b = [1 - F(S^b, \mathcal{G})]k\frac{(\#B - 2)}{\#B}. \qquad (45)$$

Since they are the shadow prices of dual constraints, it follows $\phi^b = -\eta^b$, so then (38) implies,

$$1 = 2F_S'(S^b, \mathcal{G})k\frac{(\#B - 1)}{\#B}. \qquad (46)$$

Equation (46) determines the unique optimal, $\tau_s^{\#B}$, for a given market size, $\#B$. Finally, the agent who decides on general capital investment will choose

$$1 = 2F_G'(S^b, G)k\frac{(\#B - 1)}{\#B}. \qquad (47)$$

If this agent is a non-intermediary this follows from $0 = \phi k + \phi\frac{\partial P_3^b}{\partial \tau_g} + \phi\frac{\partial \alpha}{\partial \tau_g}$; if the agent is an intermediary, it follows from the analogue to equation (38). Consider, a political manifesto that proposes a τ_g greater or less than that proposed by (46). Since, the ECT is strictly concave, as is the utility function, a tax rate exists such that the tax burden is no higher, but exchange costs are lower. That is, consumption is strictly higher. However, an equilibrium outcome has been constructed that is consistent with efficient, optimizing behavior (across intermediaries and non-intermediaries), that is nevertheless inconsistent with property 3 in the definition of equilibrium. ■

Proposition 10 *All core allocations can be supported as equilibria.*

Proof. Assume that the optimal market size is greater than unity and less than infinity. Then the above first-order conditions can be used directly to construct an equilibrium that is a competitive equilibrium (since core and competitive equilibria coincide). Since the optimal market size exists, one may construct equilibrium markets. Hence, Properties 1 and 2 of the definition of equilibrium are met. Finally, no set of agents will block this allocation since it would be unable to attain a higher level of utility than under the cooperative equilibrium. ■

D Comparative Static Results

Remark 1 shows that market size and ECT allocations increase in $\#M$. Propositions 3 and 4 demonstrates that there may exist a k^* at which two (and no more than two) plans are optimal. It remains to characterize the relationship between k and transaction costs in between the k^* at which the optimal market size changes.

Proposition 11 *For a given market size greater than one, (i) transaction costs are strictly decreasing in k; and (ii) if F_S and F_G are inelastic with respect to S and G respectively, optimal investments in ECT rise unambiguously in k.*

Proof. *Transaction costs:* Recall $\kappa = (k-y)/k = \tau_s + \tau_g + 2\left(\frac{\#M-1}{\#M}\right)(1-F)$.

Thus,

$$\begin{aligned} \frac{\partial \kappa}{\partial k} &= \frac{\partial \tau_s}{\partial k} + \frac{\partial \tau_g}{\partial k} - \theta\left[\frac{\partial F}{\partial S}\left(k\frac{\partial \tau_s}{\partial k} + \tau_s\right) + \frac{\partial F}{\partial G}\left(k\frac{\partial \tau_g}{\partial k} + \tau_g\right)\right]; \\ &= -\left(\theta k\frac{\partial F}{\partial S} - 1\right)\frac{\partial \tau_s}{\partial k} - \left(\theta k\frac{\partial F}{\partial G} - 1\right)\frac{\partial \tau_g}{\partial k} - \theta\left(\tau_s\frac{\partial F}{\partial S} + \tau_g\frac{\partial F}{\partial G}\right), \end{aligned}$$

where $\theta = 2\left(\frac{\#M-1}{\#M}\right)$. And using the optimality conditions,

$$\frac{\partial \kappa}{\partial k} = -(\tau_s + \tau_g)/k < 0. \tag{48}$$

ECT investment: Throughout, it is assumed that $F_{SG} = F_{GS} > 0$; this is consistent with the view that specific and general capital investments are positively correlated. (14) and (15) together imply,

$$\begin{aligned} \left\{F_{SS}(S,G)\left[\tau_s + \frac{\partial \tau_s}{\partial k}k\right] + F_{SG}(S,G)\left[\tau_g + \frac{\partial \tau_g}{\partial k}k\right]\right\}k + F_S(S,G) &= 0; \quad (49) \\ \left\{F_{GG}(S,G)\left[\tau_g + \frac{\partial \tau_g}{\partial k}k\right] + F_{GS}(S,G)\left[\tau_s + \frac{\partial \tau_s}{\partial k}k\right]\right\}k + F_G(S,G) &= 0. \quad (50) \end{aligned}$$

It should be recalled that we are characterizing optimal outcomes for a given $\#M$. From equation (50) it follows that $\frac{\partial \tau_g}{\partial k}k + \tau_g = -\frac{F_G}{kF_{GG}} - \frac{F_{SG}}{kF_{GG}}\left(\tau_s + \frac{\partial \tau_s}{\partial k}k\right)$, and combining with (49), one recovers an expression for the elasticity of τ_s with respect to k:

$$\frac{\partial \tau_s}{\partial k}\frac{k}{\tau_s} = \frac{F_G F_{SG} - F_S F_{GG}}{\tau_s k\left[F_{SS}F_{GG} - (F_{SG})^2\right]} - 1. \tag{51}$$

Equation (51) allows one to check the sign of the elasticity; $\tau_s k\left[F_{SS}F_{GG} - (F_{SG})^2\right]$ is positive by the assumption of strict concavity and $F_G F_{SG} - F_S F_{GG}$ is positive since $F_G, F_{SG}, F_S > 0$ and $F_{GG} < 0$. As such, for $\frac{\partial \tau_s}{\partial k}\frac{k}{\tau_s} > 0$, one requires that $F_{SG}(F_G + \tau_s k F_{SG}) > F_{GG}(F_S + \tau_s k F_{SS})$. The left hand side is positive; a sufficient condition then is that the right hand side is negative. This requires that F_S is inelastic with respect to S, i.e., that$-\frac{SF_{SS}}{F_S} < 1$. Assuming the equivalent condition holds for general capital, optimal investments in general and specific capital are, for a given market size, strictly increasing in the level of the endowment. ■

E Numerical solution to the model

First, the choice of functional form for the ECT and utility function is discussed, as is the modelling of consumption-good productivity. Next, the solution of the model is explained and analyzed holding fixed the level of the endowment. The solution of the model is then studied when k varies. Finally, the robustness of the analysis to the risk aversion and ECT parameters is examined.

E.1 A numerically tractable model

Consider a general CES form for the ECT,

$$F(S,G) = \left[\delta F_s(S)^\sigma + (1-\delta) F_g(G)^\sigma\right]^{1/\sigma},$$

where $\delta \in (0,1)$ and $1/(1-\sigma)$ is the constant elasticity of substitution. Functions F_x ($x = s$, $x = g$) are continuous, decreasing and strictly concave in their only argument; in particular, $F_x' > 0$, $F_x'' < 0$ and $F_x(0) = 0$, $F_x(y) \to 1$ as $y \to \infty$, $F_x'(y) \to 0$ as $y \to \infty$ and $F_x'(0) = \infty$ for $x \in \{s, g\}$. F_g and F_s need not be identical functions. Thus,

$$\begin{aligned} F_s &= \left[1 - \exp(-\beta \tau_s k)\right]^{\gamma_s}; \\ F_g &= \left[1 - \exp(-\beta \tau_g k)\right]^{\gamma_g}, \end{aligned}$$

where γ_s need not be equal to γ_g and where $\beta, \gamma_s, \gamma_g \in (0,1)$. These functions and parametric restrictions have a number of useful properties that both satisfy the restrictions on the ECT and facilitate numerical analysis. Taking the first partial of F_s with respect to τ_s one obtains,

$$F_s' = \beta k \gamma_s \exp(-\beta \tau_s k)\left[1 - \exp(-\beta \tau_s k)\right]^{\gamma_s - 1}$$

The Inada-type requirements are readily confirmed. In addition, the parameters $\beta, \gamma_s, \gamma_g \in (0,1)$ allow one to vary the effectiveness of the ECT.[25] one may use the parameters γ_s, γ_g to make distinctions between the relative efficiency of specific and general investments to exchange cost alleviation.

Agents have identical CRRA utility, $U(x) = \left[x^{1-\gamma} - 1\right] / (1-\gamma)$ with $\gamma > 0$ being the coefficient of relative risk aversion. For numerical tractability, we specialize to a two-state case in which production technologies are restricted to $\Lambda = \{\lambda_1, \lambda_2\}$ where $\lambda_1 < \lambda_2$ and $p(\lambda_1) = \rho$ and $p(\lambda_2) = 1 - \rho$.

The expected average technology for a market of size $\#M$ is given by the following expression

$$E\left[\bar{\lambda}_{\#M}\right] = \sum_{i=0}^{\#M} \left\{ \binom{\#M}{i} \rho^{\#M-i} (1-\rho)^i \left[(\#M - i)\lambda_1 + i\lambda_2\right] \right\} / \#M, \quad (52)$$

which is invariant to $\#M$. One may now state the problem:

$$\Gamma = \arg\max_{\#M} \sum_{i=0}^{\#M} \binom{\#M}{i} \rho^{\#M-i} (1-\rho)^i U \left[\begin{array}{c} \left\{(1 - \tau_s - \tau_g)\, k - \frac{(2\alpha)(\#M-1)}{\#M}\right\} \times \\ \times \left[(\#M - i)\lambda_1 + i\lambda_2\right] / \#M \end{array} \right]$$

subject to

$$2\left(\frac{\#M - 1}{\#M}\right) \left[\delta \left[1 - \exp\left(-\beta\tau_s k\right)\right]^{\gamma_s \sigma} + (1-\delta)\left[1 - \exp\left(-\beta\tau_g k\right)\right]^{\gamma_g \sigma}\right]^{\frac{1-\sigma}{\sigma}} \times \\ \times \delta\beta k \gamma_s \exp\left(-\beta\tau_s k\right) \left[1 - \exp\left(-\beta\tau_s k\right)\right]^{\gamma_s \sigma - 1} = 1,$$

$$2\left(\frac{\#M - 1}{\#M}\right) \left\{\delta \left[1 - \exp\left(-\beta\tau_s k\right)\right]^{\gamma_s \sigma} + (1-\delta)\left[1 - \exp\left(-\beta\tau_g k\right)\right]^{\gamma_g \sigma}\right\}^{\frac{1-\sigma}{\sigma}} \times \\ \times (1-\delta)\,\beta k \gamma_g \exp\left(-\beta\tau_g k\right) \left[1 - \exp\left(-\beta\tau_g k\right)\right]^{\gamma_g \sigma - 1} = 1,$$

where $\alpha = \left\{1 - \left[\delta\left[1 - \exp\left(-\beta\tau_s k\right)\right]^{\gamma_s\sigma} + (1-\delta)\left[1 - \exp\left(-\beta\tau_g k\right)\right]^{\gamma_g\sigma}\right]^{1/\sigma}\right\} k$ and $\max\{\Gamma\}$ is selected as the unique solution. The two constraints to solve simultaneously for the optimal $\tau^* = \left\{\tau_s^*, \tau_g^*\right\}$.

[25] We use MATLAB to compute equilibria. MATLAB code and simulation output is available from the authors.

E.2 ECT, diversification and utility with a fixed endowment

Table 2 gives the parameter values in the baseline case.

Table 2: Baseline Calibration for Endogenous Exchange Costs

endowment	k	25
coefficient of relative risk aversion	γ	3
specific ECT curvature	γ_s	0.065
general ECT curvature	γ_g	0.085
ECT factor	β	0.03
weight on specific ECT	δ	0.5
CES coefficient s, $1 - \frac{1}{s} =$	σ	0.5
low technology	λ_1	1
high technology	λ_2	5
probability of low technology	ρ	0.5

The baseline case supposes a difference in the effectiveness of S-capital and G-capital; $\gamma_g > \gamma_s$ and so G-capital is the less efficient.

First the effect on optimal market size and expected utility of different levels of ECT investment is examined. Agents choose how far (if at all) to diversify given their residual endowment and exchange costs. Figure 4 displays expected utility solving at optimal market size over a grid of $\{\tau_s, \tau_g\}$; figure 5 gives choices of market size.

Figures 4 and 5 make clear that over some combinations of the $\{\tau_s, \tau_g\}$ pairs, expected utility from diversification is higher than from not diversifying. There is a peak in expected utility at some unique combination of τ_s and τ_g. Clearly, market size does not peak at the peak of expected utility: Sub-optimal institutions can induce *larger* markets, as was explored in Section 6.

E.3 ECT, diversification and utility with a varying endowment

Using the parameter values given in Table 2, k is permitted to vary. The solution algorithm evaluates $\left\{\tau_s^*, \tau_g^*\right\}$ using the marginal conditions for market sizes up to some $\#\hat{M} > 1$ and selects the diversification level that maximizes expected utility (increasing $\#\hat{M}$ as required when $\#\hat{M}$ is the utility maximizing choice). Figure 1 gives results for these simulations.

E.4 Robustness to risk aversion and ECT parameters

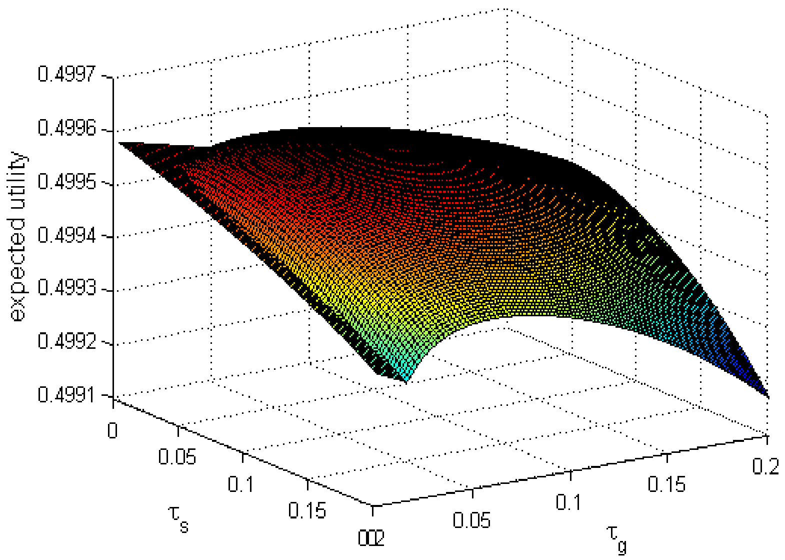
There are computational limitations to numerical analysis of the model. MATLAB v.7, for example, will only calculate up to 170!. As can be seen in the bottom-left panel of Figure 1, this will quickly become a limiting problem. However, one can see the effect of changes in some other parameters. Figures 6 and 7 give results under different parameterizations of risk aversion and ECT efficiency. In each, the central case (i.e., the middle line) reflects the parameterization in Table 2. Variations on the coefficient of relative risk aversion, γ, and on the ECT, β, are Table 2 values multiplied by $1 \pm \frac{1}{6}$.

Figure 6 shows some variation in optimal behavior in regard to risk aversion. For a given endowment, the optimal market size is decreasing in risk aversion. Further, agents are willing to spend more on diversification (i.e., forming markets), as can be seen by the size of transaction costs in each case. Figure 7 demonstrates the effect of varying the coefficient on the ECT. Reducing β means that agents diversify at lower endowments. The size of transaction costs is lower for a given endowment, but relatively unchanged for a given market size. The effect of changing the exchange technology is primarily to make it feasible for agents to diversify with a lower endowment: The nature of that diversification is little affected.

E.5 Distortive institutions and the transaction tax

For distortive institutions, the algorithm is a simple extension of that described above: One uses the marginal condition to find τ_s^* for a grid of τ_g and identify the τ_g which obtains the two rules of thumb. For the transaction tax, the value for t is introduced to the expression for α and agents optimize over investments and market size taking t as given.

Figure 4: Expected Utility and ECT Investment



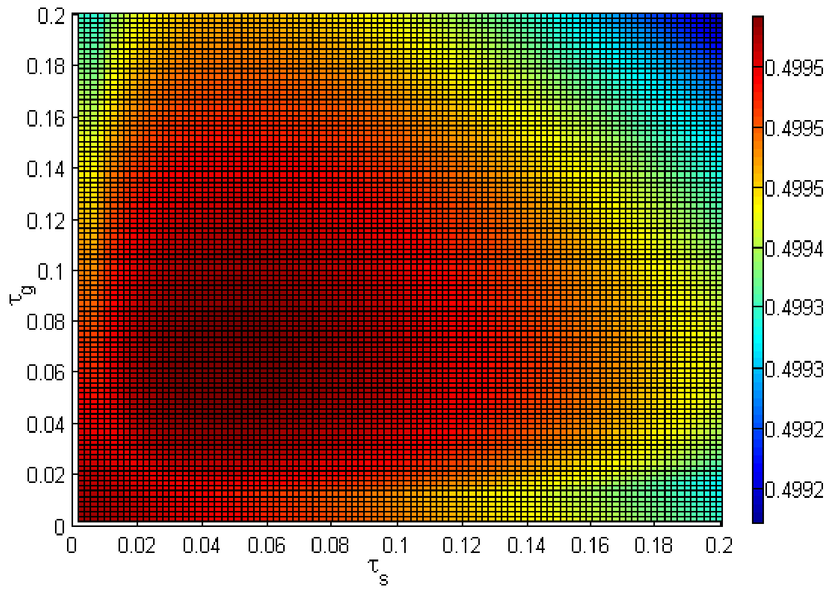


Figure 5: Optimal Market Size and ECT Investment

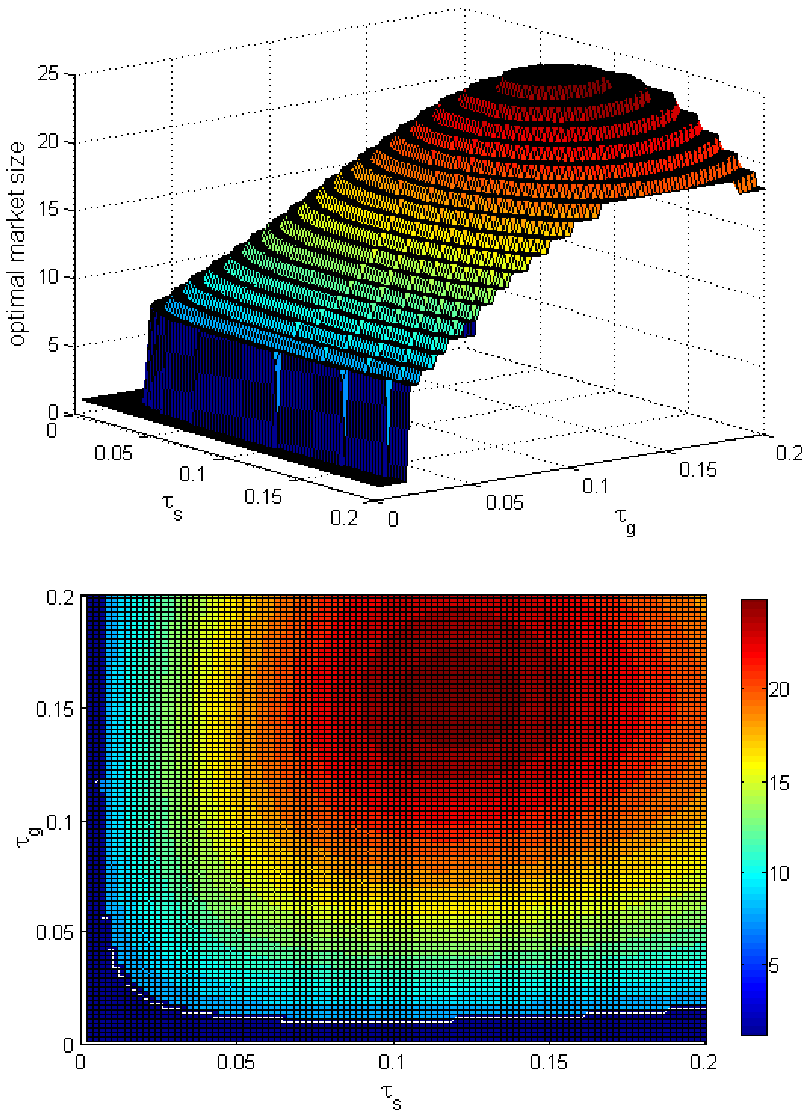


Figure 6: Risk Aversion and Endogenous Exchange

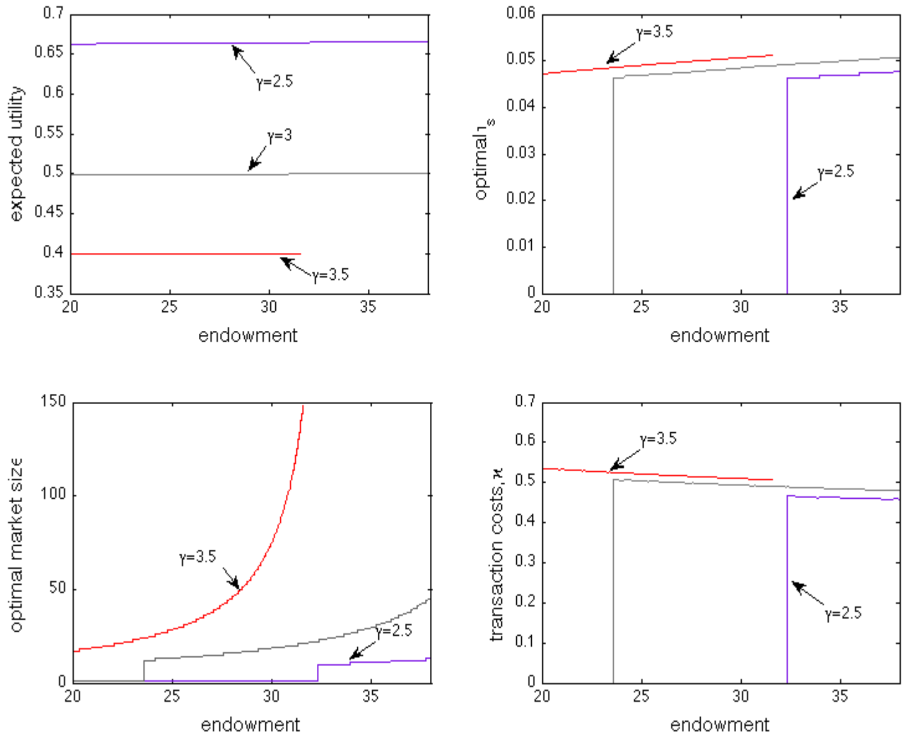


Figure 7: ECT Calibration and Endogenous Exchange