

Tax Evasion on a Social Network

Duccio Gamannossi Degl'Innocenti^{1,2} Matthew D. Rablen^{2,3}

¹Tax Administration Research Centre

²University of Exeter

³University of Sheffield



- Tax evasion causes **significant losses of public revenues** (£4.4 bn. in UK)
- Growing interest among tax authorities in **how social attitudes to tax evasion are formed**
- "Big data" information systems potentially allow tax authorities to **perceive social networks to an unprecedented degree**
- Predictive tools find **patterns in data arising due to the determinants of subjects' decisions**
- We investigate the **impact of social network on tax evasion decisions** and develop a framework to **asses the value of social network data**
 - Is it worthwhile for a tax authority to invest in this technology?

- **Standard model** of tax evasion treats it as a **private decision**
- **More recent** work allows for **social interactions** to affect compliance (Myles and Naylor, 1996 ; Hashimzade *et al.*, 2014 ; Goerke, 2013)

Limitations of Existing Literature

- Taxpayers typically assumed to know aggregate-level statistics
- Implicitly **presupposes** that the **network is the complete** network
 - but taxpayers may rely on **heterogeneous "local" information**
 - Also ruling out **heterogeneity in social connectedness**
- Other papers relax the complete network, but maintain other rigidities, i.e., **fixed pattern of connectivity, undirected network**

- The **social networks so far used** in the literature seem to **deviate importantly from real-world networks**
- We study a model allowing for an **arbitrary network**
- **Local relative consumption externalities**, heterogeneous across taxpayers
- Theoretical underpinnings to **network equilibria**

Our analysis has focused on **two** questions:

- 1 Is it possible to characterize **optimal evasion** in presence of relative utility and how do **social interactions** affect it?
- 2 How much does the **availability of more information** (especially related to social network) improves the capacity of a tax authority to **infer audit revenue effects**?

Preliminaries

- A taxpayer i has true income W_i on which they **should pay tax** $\theta(W_i)$.
- Taxpayer **may choose to evade** an amount of tax $E_i \in (0, \theta(W_i))$
- Evasion is a **risky** activity:
 - The **tax agency** is actively seeking to detect and **shut-down** evasion
 - There is a compound probability p_i that:
 - **The taxpayer is discovered** under declaring
 - **The tax agency is successful** in shutting down evasion
- The tax authority levies a **fine** $f > 1$ proportional to the evaded tax debt upon successful action
- Taxpayers care about **relative utility**
 - they benchmark consumption against a reference level R

The taxpayer's problem

$$\max_{E_i} \mathbb{E}(U_i) \equiv [1 - p_i] U(C_i^n - R_i) + p_i [U(C_i^a - R_i)]$$

After-tax income if not audited

$$C_i^n \equiv X_i + E_i$$

After-tax income if audited

$$C_i^a \equiv C_i^n - fE_i$$

Utility is linear-quadratic

$$U(z) = z[b_i - \frac{a_i z}{2}]$$

The **Privately Optimal Evasion** at an interior solution is:

$$E_i = \frac{1-p_i f}{a_i \zeta_i} \{b_i - a_i [X_i - R_i]\}$$

$$\zeta_i = [1 - p_i f]^2 + p_i [1 - p_i] f^2 > 0$$

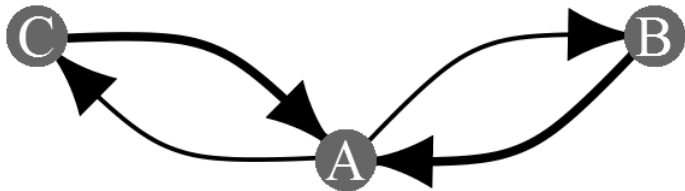
Endogenising Reference Consumption

- **Observability of consumption** summarised by a **directed network** (graph), where a link (edge) from taxpayer (node) i to taxpayer j indicates that i observes j 's consumption
- Links are **subjectively weighted**
 - some members of the reference group may be more focal comparators
- **Network** of links is represented as an $N \times N$ (adjacency) matrix, \mathbf{G} , of **subjective comparison intensity weights** $g_{ij} \in [0, 1]$,
- The weights satisfy

$$g_{ii} = 0; \quad \sum_{j \in \mathcal{R}_i} g_{ij} = 1$$

- The **set of taxpayers** whose consumption is **observed** by taxpayer i is termed i 's **reference group**, \mathcal{R}_i

An Example



$$\begin{array}{c} A \ B \ C \\ A \ \begin{pmatrix} 0 & .5 & .5 \end{pmatrix} \\ B \ \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \\ C \ \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \end{array} \equiv G$$

- Reference consumption taken as the **weighted average of expected consumption** over the members **of the taxpayer reference group** \mathcal{R}

$$R_i = \sum_{j \in \mathcal{R}_i} g_{ij} \mathbb{E}(\tilde{C}_j)$$

Where:

$$\begin{aligned} \mathbb{E}(\tilde{C}_j) &= [1 - p_j] C_j^n + p_j C_j^a \\ &= X_j + [1 - p_j f] E_j \end{aligned}$$

Nash Equilibrium – Bonacich Centrality

- **Network centrality** is a concept developed in sociology to quantify the **influence or power** of actors in a network
- **Multiple definitions:** Bonacich centrality (Bonacich, 1987) relevant in our setting

Definition

Consider a network with (weighted) adjacency matrix \mathbf{A} . For a scalar β and weight vector α , the weighted Bonacich centrality vector is given by $\mathbf{b}(\mathbf{A}, \beta, \alpha) = [\mathbf{I} - \beta\mathbf{A}]^{-1} \alpha$ provided that $[\mathbf{I} - \beta\mathbf{A}]^{-1}$ is well-defined and non-negative.

- The weighted Bonacich centrality computes the (α -weighted) sum of paths originating from a taxpayer in the network
- Longer paths are discounted by the (geometric) factor β

Proposition

If

- (i) utility is linear-quadratic, $U_i(z) = [b_i - \frac{a_i z}{2}]z$, with $a_i \in (0, \frac{b_i}{W_i})$ and $b_i > 0$ for all $i \in \mathcal{N}$;
- (ii) $1 > \rho(\mathbf{M})$; $[\mathbf{I} - \mathbf{M}] \theta(\mathbf{W}) - \alpha > \mathbf{0}$;

then there is a unique interior Nash equilibrium, at which the optimal amount of tax evaded is given by

$$\mathbf{E} = \mathbf{b}(\mathbf{M}, 1, \alpha),$$

where

$$m_{ij} = \frac{[1 - p_i f][1 - p_j f]}{\zeta_i} g_{ij};$$
$$\alpha_{i1} = \{[1 - p_i f] / [a_i \zeta_i]\} \{b_i - a_i [X_i - R_i(\mathbf{X})]\}.$$

Generalization of optimal evasion result

What **if utility is not linear-quadratic**?

For an **arbitrary** twice differentiable **utility function** considering the FO linear approximation around a Nash equilibrium to the set of FOC, it is:

$$\mathbf{E} = \mathbf{J}\mathbf{E} + \hat{\boldsymbol{\alpha}} = [\mathbf{I} - \mathbf{J}]^{-1} \hat{\boldsymbol{\alpha}}$$

Where **J** is a matrix of coefficients measuring actions' interactions

A solution is again in the form of a
weighted Bonacich centrality measure

- The model exhibits **strategic complementarities in evasion choices**
 - an increase in evasion by one taxpayer induces others to do likewise.
- Formally, **expected utility is supermodular in cross evasion choices**:

$$\frac{\partial^2 \mathbb{E}(U_i)}{\partial E_i \partial E_j} = a_i g_{ij} [1 - p_i f][1 - p_j f] > 0 \quad j \in \mathcal{R}_i$$

- How is optimal evasion impacted by information carried through the social network?

$$\frac{\partial E_i}{\partial W_j} = b_{1i} \left(\mathbf{M}, 1, \frac{\partial \alpha}{\partial X_j} \right) \geq 0;$$

$$\frac{\partial E_i}{\partial p_j} = b_{1i} \left(\mathbf{M}, 1, \frac{\partial \mathbf{M}}{\partial p_j} \mathbf{E} + \frac{\partial \alpha}{\partial p_j} \right) \leq 0.$$

- Results can be strengthened to strict inequalities if \mathbf{G} is *connected*

The Value of Network Information

- **Observing links in social networks** ought to help tax authorities to **target better** their limited **audit** resources
- Can tax authorities observe links in social networks?
 - Some individuals – celebrities – for whom it is common knowledge that many people observe them
 - “big data”
- The UK tax authority (HMRC) uses a system known as “Connect”
 - cross-checks public sector and third-party information
 - system produces “spider diagrams” linking individuals to other individuals and to legal entities such as “property addresses, companies, partnerships
- IRS also known to have also invested in big data heavily
 - but much more reticent in revealing its capabilities

Audit targeting

- Tax authority chooses **audit targets conditional** on observing each taxpayers' self-reported **income declaration** (d_i)
- By definition

$$E_i = \theta(W_i) - \theta(d_i)$$

- So

$$d_i \equiv \hat{d}_i(\mathbf{G}) = \theta^{-1}(\theta(W_i) - E_i(W_i; \mathbf{G})).$$

- We invert this function to obtain

$$W_i \equiv \hat{W}(d_i; \mathbf{G}) = \hat{d}_i^{-1}(d_i)$$

- This gives the true income W_i of a taxpayer who optimally declares an income d_i .

Limited network information

- If tax authority observes \mathbf{G} (and the remaining model parameters) it can use $\hat{d}_i^{-1}(d_i)$ to recover the true incomes

$$\hat{W}(d_i; \mathbf{G}) = W_i$$

- If the tax authority **does not perfectly observe \mathbf{G}** , but instead some (related) network \mathbf{G}' , **estimates** of the W_i **will be incorrect**

$$\hat{W}(d_i; \mathbf{G}') \neq W_i$$

- Noise in the \hat{W} feeds through into noise in the $\hat{E} = \theta(\hat{W}_i) - \theta(d_i)$
- Suppose the tax **authority observes only a subset of the links** in the network
 - $\kappa \in [0, 1]$ is the **probability** that the tax authority **observes a given link** in the social network
 - **Network observed** by the tax authority denoted $\mathbf{G}(\kappa)$ generated by randomly deleting links (with probability $1 - \kappa$)

- Audits targeted to the $100\bar{p}\%$ of taxpayers with the **highest** \hat{E}
 - Reminiscent of US “DIF score”, and similar to UK audit selection rules
- Full-information auditing gives revenue (in tax and fines)
 $\mathfrak{R}_{\max} = \mathfrak{R}(G(1))$
- No-information (random) auditing gives $\mathfrak{R}_{RA} = fpE$
- Metric used to assess value of **social network information**:

$$\Psi(\kappa) \equiv \frac{\mathfrak{R}(G(\kappa)) - \mathfrak{R}_{RA}}{\mathfrak{R}_{\max} - \mathfrak{R}_{RA}} \times 100.$$

The Social Network

- We generate a static network using the Bianconi-Barabási **fitness** model
 - Taxpayers with **higher wealth** have a higher probability of making new connections
 - Taxpayers already **heavily connected** have a higher probability of making new connections (sublinear preferential attachment, $\phi < 1$)

Formally:

$$\Pi_i = \frac{W_i [d^{in}(i)]^\phi}{\sum_{j \in \mathcal{N}} W_j [d^{in}(j)]^\phi}$$

The resulting **static** social networks used in our simulations resembles the ones observed empirically

- Tax system is linear: $\theta(W) = \theta W$
- Power law distribution of income
- Baseline parameter values
 - $\phi = 0.43$ (Pham *et al.*, 2016)
 - $N = 200$
 - $a = 2$
 - $b = 80$
 - pf calibrated to achieve evasion of 10%

Lemma

Under a linear income tax, the income of a taxpayer who declares income optimally is given by

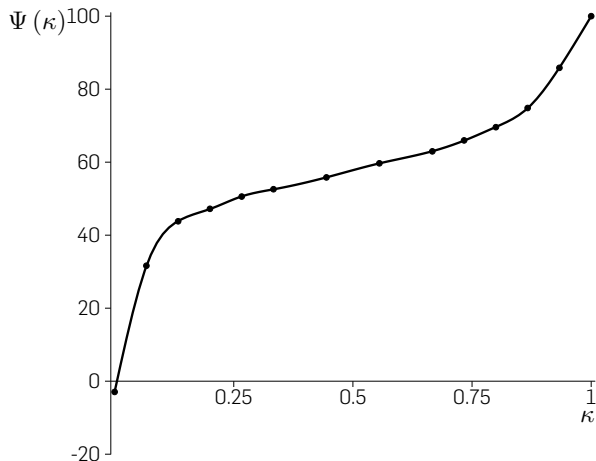
$$\hat{\mathbf{W}}(\mathbf{d}; \mathbf{G}) = \mathbf{b}(\mathbf{V}, \theta, \gamma),$$

where

$$v_{ij} = \frac{\zeta_i}{\xi_i} m_{ij}; \quad \xi_i = [1 - \theta][1 - p_i f] + \theta \{1 + [f - 2] p_i f\} > 0;$$
$$\gamma_{ij} = \frac{\{1 + [f - 2] p_i f\} \theta a_i d_i + b_i [1 - p_i f]}{a_i \xi_i}$$
$$+ \frac{[1 - p_i f] R(\mathbf{X} - \theta [1 - p_i f] \mathbf{d})}{\xi_i}.$$

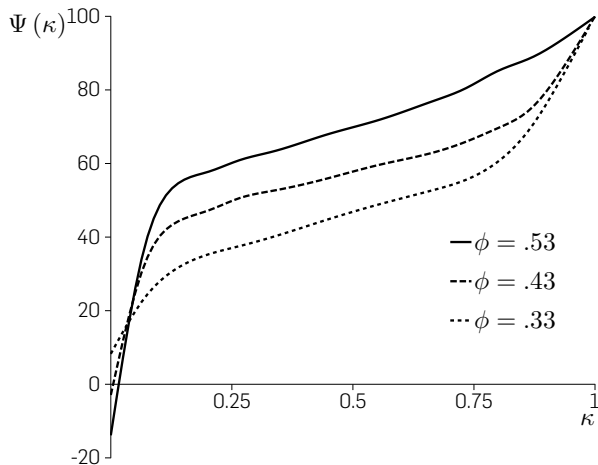
Findings – Baseline effects

- **Initial efforts** in collecting network information are characterized by **high returns**



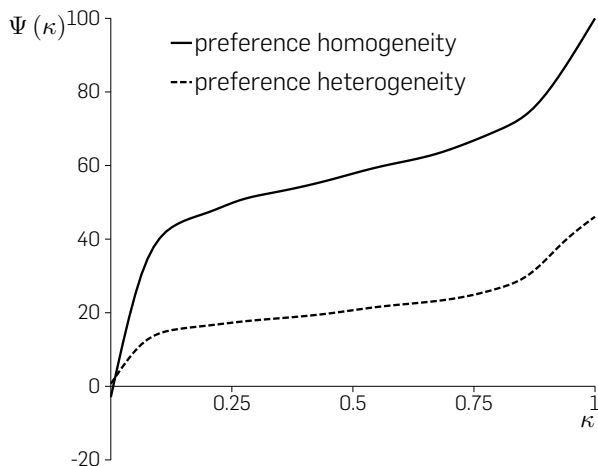
Findings – Effects of network structure

- The value of network information is **higher if preferential attachment ϕ is stronger**
- Using **predictive tools when little is known may be counterproductive** in highly concentrated networks



Findings – Effects of unobserved preference heterogeneity

- **Limited interaction** between uncertainty over **preference** and uncertainty over the **network**



Conclusions

- Our model provides a rich framework for understanding how information conveyed through a (arbitrary) social network influences optimal evasion behavior
- We show that **network information can be of value** to a tax authority
 - **strong gains to knowing a little** about the social network
 - **may actually be counterproductive** to utilize highly incomplete network information
- Some network information is **especially important in highly concentrated networks**

- Introduce **habit** (memory) dependence in reference income
 - Investigate **dynamic response** to audit interventions
 - Study **direct and indirect effects** of audit interventions
- Extend the analysis to **avoidance** and **crime** as a whole
- Analyse how adding or **removing taxpayers** from the network (detention) may affect compliance

Thank You!

Questions?