The Optimal Taxation of Network Goods

Enda Patrick Hargaden, Andrew Hanson, and Matthew Harris

August 2024

European Economic Association meetings, Rotterdam

Irish Economic Association

Seven of the ten largest companies in the world sell network goods

Long history of optimal commodity taxation literature

1. Pigou (1920) and the marginal social cost/benefit.
2. Ramsey (EJ, 1927) and the index of discouragement.
3. Sandmo (SJE, 1975) and the summation of the two.

} THE CLASSICS

4. Micheletto (JPubEc, 2008)
5. Eckerstorfer and Wendner (JPubEc, 2013)
6. Aronsson and Johansson-Stenman (AEJ:EP, 2018)
7. Farhi and Gabaix (AER, 2020)

} SOME MODERN CONTRIBUTIONS

Let's add a wrinkle to Ramsey: network effects

Consumers have their preferences over consumption x_i (at time t), $U = u(x_1^t, x_2^t, x_3^t)$ and base their decisions on the vector of final prices they face, p^t.

Governments, through taxation, can change that vector of prices. What tax structure will maximize the sum of U subject to raising enough revenue E?

$$\mathscr{L} = DWL_1 + DWL_2 + DWL_3 + \lambda\left(R_1 + R_2 + R_3 - E\right)$$

...but let's imagine x_3^t is some function of everyone else's consumption of x_3^t **(static case)** or x_3^{t-1} **(dynamic case)**.

Let's add a wrinkle to Ramsey: network effects

Consumers have their preferences over consumption x_i (at time t), $U = u(x_1^t, x_2^t, x_3^t)$ and base their decisions on the vector of final prices they face, p^t.

Governments, through taxation, can change that vector of prices. What tax structure will maximize the sum of U subject to raising enough revenue E?

$$\mathscr{L} = DWL_1 + DWL_2 + DWL_3 + \lambda\left(R_1 + R_2 + R_3 - E\right)$$

...but let's imagine x_3^t is some function of everyone else's consumption of x_3^t **(static case)** or x_3^{t-1} **(dynamic case)**.

The plan of attack

1. We derive the optimal tax rate τ for a network good in a static model. $\rightarrow$ **A positive network effect lowers the tax rate.**

2. We derive the optimal tax rates (τ_1, τ_2) for a network good in a simple (*haha!*) two period model. $\rightarrow$ **Results are complex, but $\tau_1 \neq \tau_2$**

3. We simulate the optimal tax sequence for a discrete network good in a **fuller quantitative model**, running robustness checks (e.g. changing the supply side). $\rightarrow$ **Results are complex, depend on parameter values.**

The plan of attack

1. We derive the optimal tax rate τ for a network good in a static model. $\rightarrow$ **A positive network effect lowers the tax rate.**

2. We derive the optimal tax rates (τ_1, τ_2) for a network good in a simple (*haha!*) two period model. $\rightarrow$ **Results are complex, but $\tau_1 \neq \tau_2$**

3. We simulate the optimal tax sequence for a discrete network good in a **fuller quantitative model**, running robustness checks (e.g. changing the supply side). $\rightarrow$ **Results are complex, depend on parameter values.**

Static Model

Theoretical setup

What happens if the externality affects the willingness-to-pay?

Ramsey (1927)	$u^i = u(1 - x_0, x_1, \ldots, x_m)$
Sandmo (1975)	$u^i = u(1 - x_0, x_1, \ldots, x_m, \phi)$
This paper	$u^i = u\left(1 - x_0, x_1, \ldots, x_m(\phi), \phi\right)$

In these models, x_0 is working and $1 - x_0$ is leisure. We cannot tax leisure, so it reminds an outside option for the consumer.

The solution

Sandmo's result:

$$\tau_m = (1-\mu)\underbrace{\left[\frac{-1}{P_m}\frac{\sum_{i=1}^{m} x_i J_{im}}{J}\right]}_{\textbf{Ramsey-esque}} - \mu \underbrace{\left[n\left(\frac{u_{m+1}}{u_m}\right)\right]}_{\textbf{Pigou-esque}}$$

This paper's static result:

$$\tau_m = (1-\mu)\underbrace{\left[\frac{-1}{P_m}\frac{\sum_{i=1}^{m} x_i J_{im}}{J}\right]}_{\textbf{Ramsey-esque}} - \mu \underbrace{\left[n\left(\frac{u_{m+1}}{u_m}\right)\right]}_{\textbf{Pigou-esque}} - \mu \underbrace{\left[n\left(\frac{\partial x_m}{\partial \phi}\right)\right]}_{\textbf{Network effect}}$$

These results are implicit: we are assuming interior solutions/necessary conditions here.

The solution

Sandmo's result:

$$\tau_m = (1-\mu)\underbrace{\left[\frac{-1}{P_m}\frac{\sum_{i=1}^{m} x_i J_{im}}{J}\right]}_{\textbf{Ramsey-esque}} - \mu \underbrace{\left[n\left(\frac{u_{m+1}}{u_m}\right)\right]}_{\textbf{Pigou-esque}}$$

This paper's static result:

$$\tau_m = (1-\mu)\underbrace{\left[\frac{-1}{P_m}\frac{\sum_{i=1}^{m} x_i J_{im}}{J}\right]}_{\textbf{Ramsey-esque}} - \mu \underbrace{\left[n\left(\frac{u_{m+1}}{u_m}\right)\right]}_{\textbf{Pigou-esque}} - \mu \underbrace{\left[n\left(\frac{\partial x_m}{\partial \phi}\right)\right]}_{\textbf{Network effect}}$$

These results are implicit: we are assuming interior solutions/necessary conditions here.

So what does that all mean?

1. The basic Ramsey logic still holds — tax inelastic goods. The basic Sandmo logic still holds — address the atmospheric externality...
2. ...but now we have two channels for the externality to affect welfare. One traditional/atmospheric, and one capturing the spillover into demand.

The best example here might be COVID and indoor dining. A large amount of COVID (itself a negative externality) reduces the demand for indoor dining.

The Sandmo result is **generalized** to the case where the externality affects the utility of consumption. Depending on the magnitude of the effects, the optimal tax rate can be **positive or negative**.

Two Period Model

The two period model introduces a durable good

Shut down the atmospheric channel for now. A representative household lives for two periods.

- They solve their optimization problem for a durable good c_t based on the price and network size $f(c, X)$.
- Preferences measured by ρ

The **government** observes the consumer demand functions, and commits to a tax schedule.

- Gov't endogenizes the network effect: $f(c, X) \neq \mathbf{g(c)}$.
- Chooses τ_1 and τ_2 to maximize welfare subject to raising enough revenue.

The following elements of notation are worth remembering

- The function f is the private utility of the network good, given total consumption
- The government embeds the external effect into its considerations g, so that $f \neq g$
- Define $\Lambda = \beta \frac{\lambda_2}{\lambda_1}$, a (stochastic) discount factor
- Finally, the big one, $H\left(c_1, X_1\right) \equiv \left[\frac{g(c_1)}{f(c_1, X_1)}\right]^{\rho-1} \left(\frac{\partial g(c_1)/\partial c_1}{\partial f(c_1, X_1)/\partial c_1}\right)$

This last element is a measure to which the externalities provide a non-pecuniary benefit. Does anyone have a name for this?

The optimal tax rates are solved using the IFT and Cramer's Rule

For completeness, here are the welfare-maximizing tax rates:

$$\tau_1^* = (\Lambda - 1)\,\frac{\Lambda\left[H_1 - 1\right]H_2\frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \Lambda\left[H_2 - 1\right]\frac{\partial c_2}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \left[H_1 - 1\right]\left[\Lambda\left[H_2 - 1\right] + 1\right]\frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}{\Lambda H_1 H_2\frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \left[H_1 - 1 + \Lambda\right]\left[\Lambda\left[H_2 - 1\right] + 1\right]\frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}$$

$$\tau_2^* = \frac{-\Lambda H_1\left[H_2 - 1\right]\frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} + \Lambda\left[H_1 - 1 + \Lambda\right]\left[H_2 - 1\right]\frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2} - (\Lambda - 1)^2\left(H_1 - 1\right)\frac{\partial c_1}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}{\Lambda H_1 H_2\frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \left[H_1 - 1 + \Lambda\right]\left[\Lambda\left[H_2 - 1\right] + 1\right]\frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}$$

There isn't much to glean from these without imposing structure, e.g. making assumptions relative magnitudes of income versus substitution effects. But we can say that, in general, $\tau_1^* \neq \tau_2^*$.

The optimal tax rates are solved using the IFT and Cramer's Rule

For completeness, here are the welfare-maximizing tax rates:

$$\tau_1^* = (\Lambda - 1)\, \frac{\Lambda\left[H_1 - 1\right] H_2 \frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \Lambda\left[H_2 - 1\right] \frac{\partial c_2}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \left[H_1 - 1\right]\left[\Lambda\left[H_2 - 1\right] + 1\right] \frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}{\Lambda H_1 H_2 \frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \left[H_1 - 1 + \Lambda\right]\left[\Lambda\left[H_2 - 1\right] + 1\right] \frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}$$

$$\tau_2^* = \frac{-\Lambda H_1\left[H_2 - 1\right] \frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} + \Lambda\left[H_1 - 1 + \Lambda\right]\left[H_2 - 1\right] \frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2} - (\Lambda - 1)^2 (H_1 - 1) \frac{\partial c_1}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}{\Lambda H_1 H_2 \frac{\partial c_1}{\partial \tau_1}\frac{\partial c_2}{\partial \tau_2} - \left[H_1 - 1 + \Lambda\right]\left[\Lambda\left[H_2 - 1\right] + 1\right] \frac{\partial c_2}{\partial \tau_1}\frac{\partial c_1}{\partial \tau_2}}$$

There isn't much to glean from these without imposing structure, e.g. making assumptions relative magnitudes of income versus substitution effects. But we can say that, in general, $\tau_1^* \neq \tau_2^*$.

The paper also provides sufficient conditions for $\tau_1^* < \tau_2^*$

We provide one numerical example of what tax rates look like. Suppose:

$$f(c, X) = \gamma c + (1/2)\alpha X^2$$
$$g(c) = \gamma c + (1/2)\alpha c^2$$

with parameter values $\alpha = 0.5, \beta = 0.99, \theta = 0.8, \rho = 0.9, \gamma = 5$. Then the optimal strategy is a subsidy of 17.7 percent followed by a tax rate of 204 percent.

But then: so what?

The formulas are messy, the assumptions are restrictive, the equilibria may not be unique, and the parameter values are highly sensitive to specification.

Is this stuff in any way robust?

Multi-period quantitative model

The basic setup

- There are six periods in the model. Consumption is initially zero.
- Consumers decide whether to buy this good or not. The up-front price is p and it comes with a per period tax rate τ.
- Once the individual has bought the good, they do not 're-buy' the good in each successive period — they just pay a usage tax.
- Our 'exogenous' revenue constraint R is internally determined: collected revenue with a static tax rate of either 0.3 or 0.5. We calculate tax revenue and consumer surplus in each of these scenarios.

Consumers use information available to them at time t

Following Goyal (2012), consumers' expectation of state variables (such as the size of the network) in time $t+1$ is their contemporaneous value, i.e. $\mathbb{E}_t\left[x_{t+1}\right] = x_t$.

Consumers are not perfectly forecasting the future. In particular, they do not incorporate the benefits of future growth.

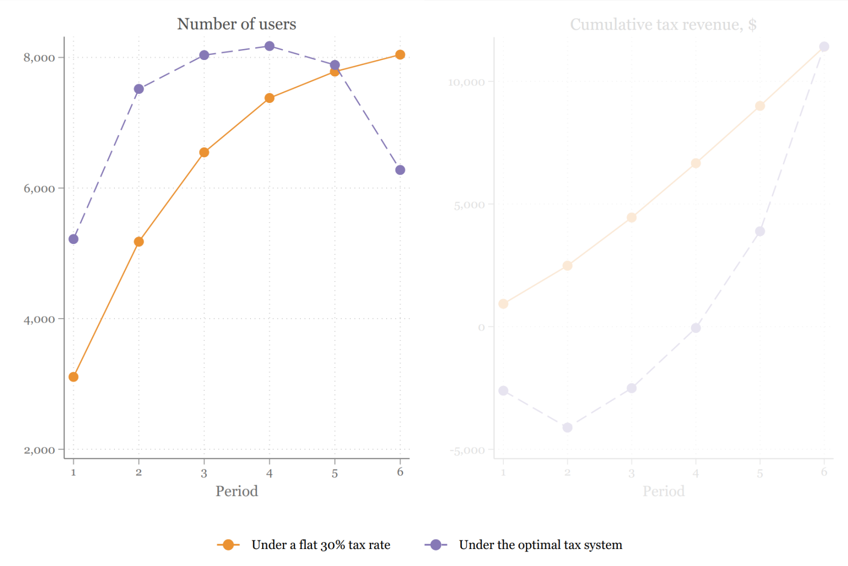
Intertemporal Fiscal Externality

Note: This is one way to model an intertemporal fiscal externality. Encouraging network growth today affects the ability to raise taxes next period.

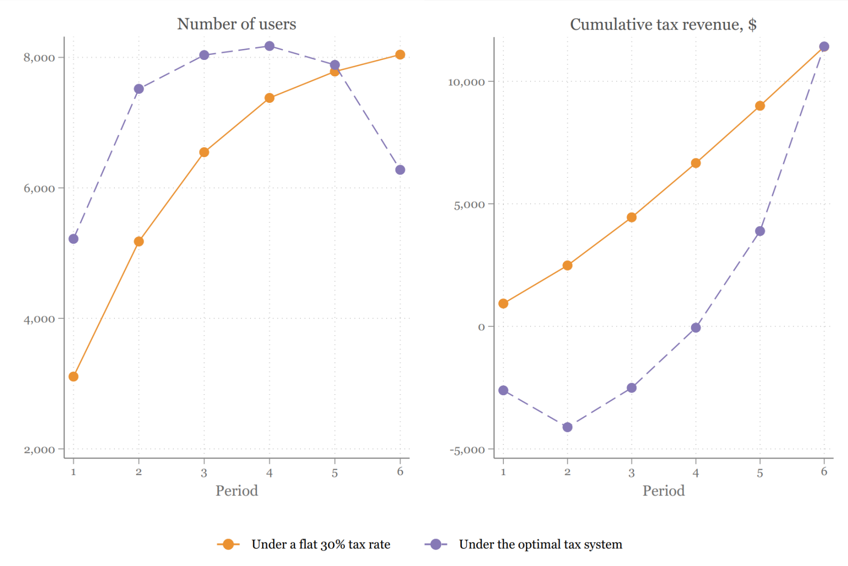
Additional features of the model

1. Consumers may freely dispose of the good at any point.
2. Supply side is competitive. Adjusting this to a price-setting firm doesn't substantially change the results.
3. In general, we assume the network good creates negative atmospheric externalities (δ). Consider Bitcoin, Airbnb, Waze. But we also let δ be zero or positive.
4. Government grid-searches over all possible tax rates (usually $[0, 1]$ in ten steps), and chooses the one that maximizes welfare subject to raising revenue greater than that raised by a time-invariant sequence.

One depiction of time-invariant rates versus optimal tax rates



One depiction of time-invariant rates versus optimal tax rates



Conclusions

Conclusions

Network goods are an increasingly important part of the economy, and a growing share of the tax base.

- Borrowing insights from the dynamic IO pricing literature, the optimal tax sequence on a network good is not constant through time.
- The variance of the optimal tax rate is increasing in the strength of the network effect, consistent with infant industry-type arguments.
- The model is likely applicable to any number of goods with subtle network externalities, e.g. indoor dining during a pandemic.

The Good News: in a world with uncertainty, a wait-and-see policy might not be a bad option.

Thank you for your time