Subsidies

Prof. Richard Sweeney

ECON8852, Boston College

Can we use subsidies instead of Pigouvian taxes?

Tax on emissions:

$$\pi(x, e) = px - C(x, e) - \tau e$$
$$p = C_x(x, e)$$
$$-C_e(x, e) = \tau$$

Can we use subsidies instead of Pigouvian taxes?

Tax on emissions:

$$\pi(x, e) = px - C(x, e) - \tau e$$
$$p = C_x(x, e)$$
$$-C_e(x, e) = \tau$$

Subsidy on emissions reductions below baseline ($\bar{e}$):

$$\pi(x, e) = px - C(x, e) + s(\bar{e} - e)$$
$$-C_e(x, e) = s$$

On first glance, it looks like the are equally efficient

The downside of subsidies

Subsidy equivalent to a tax plus a lump sum transfer

$$\pi(x, e) = px - C(x, e) + s(\bar{e}) - s(e)$$

Baseline emissions are unknown.

- Bigger problem is this is effectively a subsidy for operating
  - rent $s * \bar{e}$
- This in turn leads to inefficient entry/ exit.
  - [I think this is an interesting topic in electricity markets]
- Finally: no (strong) double dividend
  - need to *raise* tax revenue $s(\bar{e} - e)$
  - rather than *generating* revenue τe
  - Note that "strong" DD unlikely

What about subsidizing a substitute?

- Two methods for producing a good: green (G) or dirty (D)
  - convex costs so both supplied
- Product is homogenous and competitively supplied at price P

$$Q(P) = Q^G(P) + Q^D(P)$$

- Dirty production is associated with per unit externality (τ)

Pigouvian tax

- Know the quantity of dirty good goes down
  - $Q^D(P' - \tau) < Q^D(P)$
- Know the price must go up
  - $P' > P$ implies $Q' < Q$
- Implies that green good goes up, but does not fully offset
  - $0 < \Delta Q^G < |\Delta Q|$

Pigouvian *subsidy* on green good

– subsidy must mean that production of green good increases

  – $Q^G(P' + \tau) > Q^G(P)$

– So price must *fall*

  – Q < Q' implies P > P'

– Dirty good falls, but net supply increases

  – $0 < |\Delta Q^G| < \Delta Q$

now DWL *must* be positive, since P < social MC

Things look even worse if the cost of public funds is positive

  – subsidy must be financed with distortionary taxes

Despite this, subsidies remain remarkably popular

– so worth understanding how they work

– focus on case of correcting externalities, but implications extend to general goal of increasing socially desirable activites

Core tax theory result: Targeting

From Kopczuk (EL 2003)

> *One of the main results of the literature on optimal taxation in the presence of externalities (e.g., Sandmo, 1975; Ng, 1980; Bovenberg and van der Ploeg, 1994) is the "additivity property": the presence of an externality affects only the formula for the tax on the externality generating commodity, and it does not affect the optimal taxes imposed on other goods.* ***Dixit (1985) referred to this result as an example of a more general "principle of targeting": one should correct the externality by targeting its source directly.*** *This result has powerful implications for designing tax policy, because it implies that correcting externalities may be done on an ad hoc basis by taxing suspect commodities, without relying on other components of the tax system to address this problem.*

Inputs vs outputs

Should subsidies pay for inputs or outputs?

- Agriculture
  - subsidize fertilizer or crops
- R & D
  - pay for scientists or innovation
- Housing
  - build structures or subsidize months rent
- Stimulus
  - investment or output

Typically policy objective tied to **output**, so informal logic suggests output subsidies are best [Schmalensee, 1980].

What margin minimizes public expenditure?

Parish and McLaren [1982] show that this informal logic is not correct, if the goal is simply to minimize government expenditure (*cost-effectiveness*)

Suppose government wants to obtain some change in output dQ.

Subsidies work *at the margin*
– by driving a wedge between MR and MC

But they payout over all **inframarginal** units as well.

Implication: Cost-effectiveness depends on whether inputs are more or less intensive at the margin.

Consider the case of a single input x

With a per unit output subsidy s, FOC

$$P + s + P'q = mc$$

If firm uses x' inputs to produce the marginal unit, the same increase in quantity could be achieved with subsidy $w = s/x'$

$$P + P'q = mc - wx'$$

The total public expenditure E would be:

- $E^O = s * q$
- $E^I = w * x$

Consider case of decreasing returns

So x is used more intensively on the margin than on average

$$x' > \bar{x} = x/q$$

Plugging into the total expenditure function:

$$E^O = sq = (wx')q > (w\bar{x})q = wx = E^I$$

So in the case of **decreasing** returns, an input subsidy would achieve $dQ(s)$ at lower public cost than an output subsidy

– opposite holds with IRS

– equivalent with CRS

What if there are multiple inputs?

Is it better to target one input or have a uniform input subsidy?

– If the inputs are not substitutable, then the previous logic applies.

– Subsidizing inputs that are used relatively more intensely at the margin is cheaper than a uniform input subsidy.

– If the inputs are substitutable, changing relative input costs will alter the input mix.

– PM show that this input substitution could potentially more than fully offset the intended effect.

PM provide a rationale for seemingly misguided input subsidies

- These abound: fertilizer, capital; transportation; services
- Conclude that a good input for a tax case would:
  - have a high elasticity of supply
  - be substitutable for factors that are relative fixed/ inelastic
  - be less substitutable with inputs that are also very elastic

Important caveat: Input subsidies may also cause allocative inefficiency

– Absent subsidies, inputs used such that marginal product is proportional to marginal price.

– By distorting relative prices, we also distort marginal products

 – Marginal subsidized input less productive than social cost

– [Goolsbee, 2004] looks several industries and finds that firms shift towards higher quality capital when they receive investment subsidies

 – this explains essentially ALL of the increase in spending, rather than buying more of existing capital types

Empirical evidence

- Cash for coolers [Davis et al., 2014]: most recipients inframarginal

- [Goolsbee, 1998]: shocks to R&D spending increase scientist and engineer salaries by 30-40%

- Not much empirical evidence on inputs vs outputs

Aldy, Gerarden & Sweeney (2021)

- Wind capacity in the US increased tenfold in the last decade
- This has largely been driven by generous Federal subsidies
  - 2014 one-year PTC extension: $6.4 billion
- 2009 ARRA introduced a new subsidy type that targeted wind investment rather than electricity output
  - Section 1603 grant program
- AGS estimate the impact of this subsidy choice on wind production

Research Question

Are wind farms less productive if they receive an investment subsidy rather than an output subsidy?

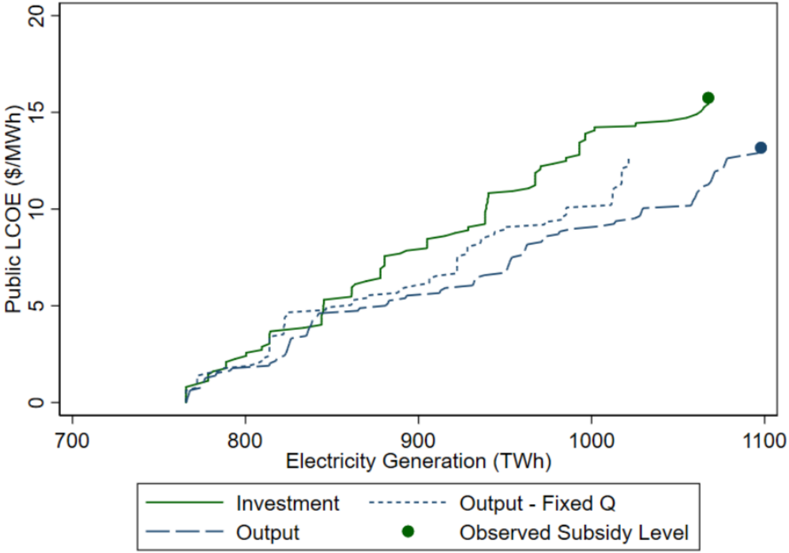
Methods:

- Use natural experiment introduced by ARRA
- Instrumental variables approach (fuzzy RD)
- Matching + "difference-in-differences"

Results:

- Wind farms selecting the investment subsidy are 10 - 12 % less productive than they would have been (under the PTC)

AGS Public Cost of Energy



Subsidy Timing

How do consumers tradeoff upfront and future energy costs/ benefits?

- One important difference between input and output subsidies is timing. Investment subsidies are up front, whereas production payments inherently occur over time.
- It's possible that adopters may value the two differently. Why?

How do consumers tradeoff upfront and future energy costs/ benefits?

- One important difference between input and output subsidies is timing. Investment subsidies are up front, whereas production payments inherently occur over time.
- It's possible that adopters may value the two differently. Why?
  - Pure time preference
  - Uncertainty
- De Groote & Verboven attempt to estimate both the discount rate and preference for subsidy type conditional on that.
  - Possible because Belgium subsidies both investment and production, and the relative generosity of each varied over time (as did prices)

Specifying preferences for upfront vs future energy costs/ benefits?

Utility to i from choosing good j:

$$u_{ij} = \nu_{ijt} + \epsilon_{ijt}$$

, where ϵ is the RUM component.

ν is the "conditional" value choosing choice j

$$\nu_{ijt} = \delta_{jt} = x_{jt}\gamma - \alpha p_{jt} + \xi_{jt}$$

"price" here consists of an upfront cost (net of up front subsidies (τ)) and a discounted stream of future subsidies s. [Will come back to this]

$$p_{jt} = p_{jt}^{GROSS} - \beta^{\tau}\tau - \beta^{s} s$$

Static estimatation

Berry (1994) showed how we can allow for unobserved ξ by inverting the observed shares:

$$S_{jt} = \frac{exp(\delta_{jt})}{\sum_0^J exp(\delta_{jt})}$$

where δ_{jt} is the mean utility, before ϵ is observed (and ignoring any heterogeneity in preferences)

Allcott & Wozny discussed how to estimate using exogenous variation in s and product fixed effects to remove ξ

What happens if consumers chose *when* to buy?

Assume that purchasing is a terminating activity, so if $d_t = j > 0$, utility is 0 for all subsequent periods.

However, if $d_t = 0$, the consumer has the option to buy again next period. So

$$u_{0t} = 0 + \beta E_t \bar{V}_{t+1}$$

where $E_t \bar{V}_{t+1}$ is the expected value of subsequent optimal behavior.

Ignoring this *continuation value* will lead to biased estimates

The bias is context specific, but several papers have found this leads agents to appear more price elastic (for example in the case of laundry detergent (Hendel and Nevo))

To use the RUM framework on dynamic decisions, we need to incorporate continuation values

What's in $E_t \bar{V}_{t+1}$?

[From Arcidiacono's nice Annual Review article.]

- d_t is the choice from J alternatives each period
- decision rule

$$\sigma^* = \arg\max_{\sigma} E\left(\sum_t^T \beta^t [u(x_t, d_t) + \epsilon(d_t)] | x_1, \epsilon_1\right)$$

  where the expectations are taken over the future realizations of x and ϵ induced by σ^*
- Value function $V(x_t, \epsilon_t)$ represents the expected present discounted value of lifetime utility from following σ^*
- Since ϵ is unobserved, its helpful to define the ex-ante value function as the value of being in state x_t, just before ϵ_t is revealed.

How to compute $E_t \bar{V}_{t+1}$?

Rust (1987) computes EV directly within estimation routine. Assumes:

1. Exogenous state variables follow a first order Markov process
2. Evolution is conditionally independent of the taste shocks ϵ

Estimation:

1. Guess θ
2. Iterate on Bellman Operator using observed state transition probabilities
3. Compute likelihood of observed decisions

Several papers have estimated solar PV adoption models this way (Burr 2016, for example)

Challenges with Rust's brute force method

- Rust's approach is computationally burdensome.

- It also requires explicit assumptions about the stochastic processes governing state variables (and agents expectations over them).
  - DG & V describe their states as $\omega_t = (u_{0,t+1}, \delta_{1,t+1}, \cdots, \delta_{J,t+1})$
  - note each δ depends on complicated policy / subsidy variation
  - In Rust we assume ω follows a Markov process matched to empirical density $f(\omega_{t+1}|\omega_t)$

As an alternative, DG & V adopt a "rational expectations" approach from Scott (2013)

- Decompose the ex ante value function into a realization and a short run prediction error

$$\eta_t \equiv \bar{V}_{t+1} - E_t \bar{V}_{t+1}$$

- Assume that households have rational expectations in the sense that η is mean zero. (will come back to this)

- This allows them to rewrite the utility of the outside option as

$$\delta_{0t} = u_{0,t} + \beta(\bar{V}_{t+1} - \eta_t)$$

De Groote and Verboven approach I

The ex ante value function can be rewritten as

$$\bar{V}_{t+1} = 0.577 + \ln \sum_{j=0}^{J} exp(\delta_{j,t+1})$$

Hotz and Miller (1993) insight: pick an option $j = 1$,

$$s_{1,t+1}(\delta_{t+1}) = \frac{exp(\delta_{1,t+1})}{\sum_0^J exp(\delta_{j,t+1})}$$

then take logs

$$\ln \sum_0^J exp(\delta_{j,t+1}) = \delta_{1,t+1} - \ln s_{1,t+1}(\delta_{t+1})$$

De Groote and Verboven approach II

Plug that back into the ex ante value function

$$\bar{V}_{t+1} = 0.577 + \delta_{1,t+1} - \ln s_{1,t+1}(\delta_{t+1})$$

What does this say? The ex ante value function at time $t+1$ is equal to

- the utility from choosing option 1
- plus Euler's constant (mean of the Type I EV distro)
- minus some correction for the chance that 1 might not be the best option (which is (intuitively?) related to the share of 1)

De Groote and Verboven approach III

Finally, we have the mean utility from not adopting as

$$u_{0t} = 0 + \beta(0.577 + \delta_{1,t+1} - \ln S_{1,t+1} - \eta_t)$$

where we have normalized the static component of the outside option to 0, and plugged in the observed share.

Back to Berry (1994)

$$In(S_{jt}/S_{0t}) = \delta_{jt} - \delta_{0t}$$

Sub in mean utilities to obtain

$$\ln(S_{j,t}/S_{0,t}) = (x_{jt} - \beta x_{1,t+1})\gamma - \alpha(p_{jt} - \beta p_{1,t+1}) + \beta InS_{1,t+1} + e_{jt}$$

Where $e_{jt} = \xi_{jt} - \beta(\xi_{1,t+1} - \eta_t)$

[in static case, $\beta = 0$]

DG & V Intuition?

Imagine there is a single product ($J = 1$)

Estimating equation can then be written

$$\ln \frac{S_{1,t}/S_{1,t+1}}{S_{0t}} = (x_{1t} - \beta x_{1,t+1})\gamma - \alpha(p_{jt} - \beta p_{1,t+1}) + e_{1t}$$

– Basically regress the change in the number of adopters on the change in price.
– Same logic as an Euler equation, where agent is indifferent between a one period delay in an action along equilibrium path.

Investment subsidies (up front)

- 2006-2007 households can apply for a 10% investment subsidy
- There was also a general tax credit of 40%
  - Max amount varied from €1,200 in 2006 to €3,600 in 2011.
- In 2012 the tax credits are abolished

Investment subsidies (up front)

- 2006-2007 households can apply for a 10% investment subsidy
- There was also a general tax credit of 40%
  - Max amount varied from €1,200 in 2006 to €3,600 in 2011.
- In 2012 the tax credits are abolished
- What do people think about this? What might consumer expectations look like?

Production subsidies (in the future)

GCC production subsidies

- Fix price per unit of output, paid over 20 years
- Level of payment varies a LOT.
  - €450 in 2006; €90 in 2012.
  - 2013 phaseout begins; 2014 it is abolished.

Net metering

- Compensated for electricity supplied back to the grid.
- Can't go negative.
- Not much said about compensation rates.

Data

What do they observe?

Data

What do they observe?

- Date and size of system
- Location

What don't they observe?

Data

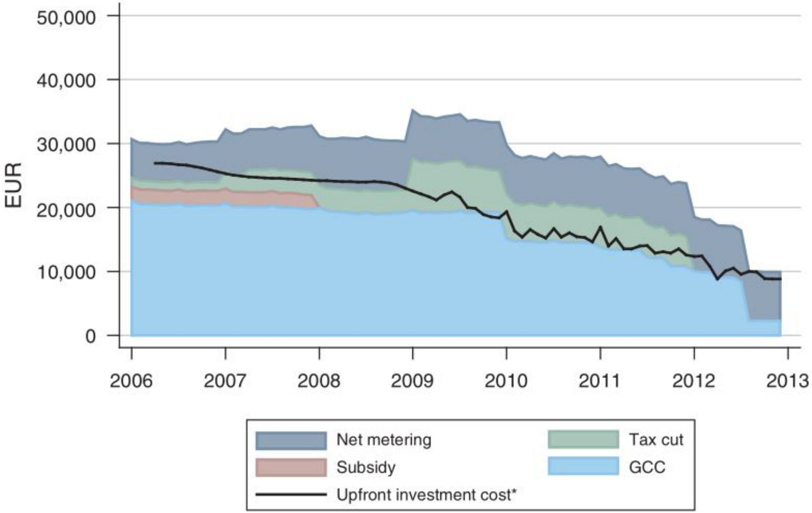
What do they observe?

- Date and size of system
- Location

What don't they observe?

- Prices
- Actual production

DG & V Variation



Which of these are actually observed well in the data?

DG & V Adoptions

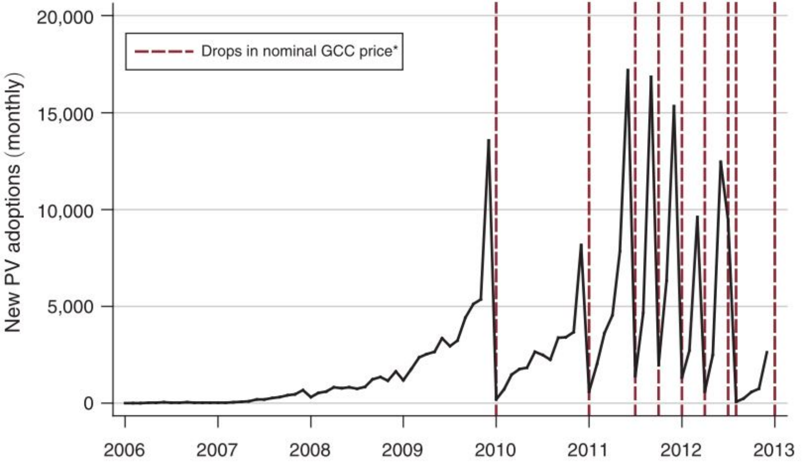


FIGURE 2. 2006–2012: TIME SERIES OF NEW PV ADOPTIONS AND DROPS IN NOMINAL GCC PRICE

Main estimates at the MONTH level.

Clearly it would be hard to express expectations for this pattern.

DG & V Results

TABLE 2—EMPIRICAL RESULTS

	Static (1)		Dynamic (2)		+ Micro-moments (3)	
Price sensitivity in 10^3 EUR ($-\alpha$)	−0.318	(0.074)	−0.470	(0.098)	−0.604	(0.100)
Monthly discount factor (β)	0.9886	(0.0016)	0.9884	(0.0025)	0.9884	(0.0024)
Annual interest rate in % ($r = \beta^{-12} - 1$)	14.82	(2.28)	15.09	(3.43)	15.00	(3.42)

- Dynamics matters a lot for price sensitivity
- Has essentially no effect on the discount rate
- After all this, funny that they come up with basically the same r as Allcott and Wozny

Main story is remarkably robust

TABLE 3—ROBUSTNESS: SPECIFICATION CHOICES

	All terminal choices		10% potential market		Time controls	
Price sensitivity in 10^3 EUR ($-\alpha$)	−0.422	(0.046)	−0.471	(0.098)	−0.439	(0.117)
Monthly discount factor (β)	0.9873	(0.0007)	0.9883	(0.0025)	0.9895	(0.0016)
Annual interest rate in % ($r = \beta^{-12} - 1$)	16.63	(1.03)	15.13	(3.44)	13.52	(2.17)
Control variables (γ)						
Alternative-specific constant						
Common constant	−10.152	(11.278)	1.557	(16.345)	−610.0	(1,017.3)
2kW	−2.045	(0.129)	−1.834	(0.562)	−1.614	(0.421)
6kW	−0.282	(0.136)	−0.507	(0.595)	−0.721	(0.460)
8kW	−2.021	(0.262)	−2.442	(1.158)	−2.879	(0.881)
10kW	−1.989	(0.399)	−2.587	(1.683)	−3.250	(1.260)
Time controls						
Linear trend					1.172	(1.985)
Spring					−0.177	(0.470)
Summer					−0.047	(0.493)
Fall					−0.021	(0.358)
Hansen's J (p-value)	31.736 ($p = 0.2854$)		Exactly identified		Exactly identified	
Observed macro moments	220		220		220	

Notes: Standard errors clustered within 44 time periods. Instruments are approximations of optimal instruments (Chamberlain 1987). Standard errors of r obtained via delta method.

Even the distribution they recover is tight

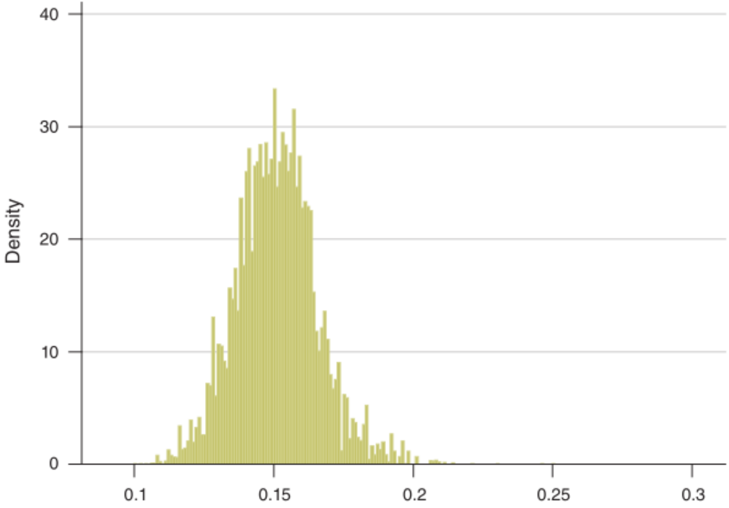
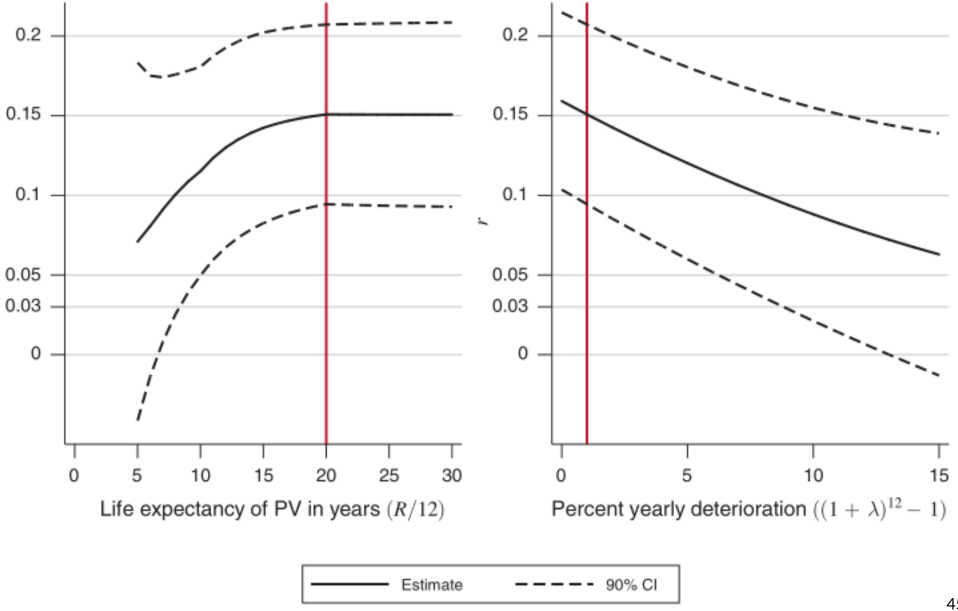


FIGURE 5. HETEROGENEOUS INTEREST RATE

Graphs on alternative assumptions



[ ]Lucas W Davis, Alan Fuchs, and Paul Gertler. Cash for coolers: Evaluating a Large-Scale appliance replacement program in mexico. *American Economic Journal: Economic Policy*, 6(4):207–238, 2014. ISSN 1945-7731, 1945-774X. doi: 10.1257/pol.6.4.207. URL http://pubs.aeaweb.org/doi/abs/10.1257/pol.6.4.207.

[ ]Austan Goolsbee. Does government R&D policy mainly benefit scientists and engineers? *The American economic review*, 88(2): 298–302, 1998. ISSN 0002-8282. URL http://www.jstor.org/stable/116937.

[ ]Austan Goolsbee. Taxes and the quality of capital. *Journal of public economics*, 88(3-4):519–543, March 2004. ISSN 0047-2727. doi: 10.1016/S0047-2727(02)00190-1. URL http://www.sciencedirect.com/science/article/pii/S0047272702001901.

[]Ross M Parish and Keith Robert McLaren. Relative Cost-Effectiveness of input and output subsidies. *Australian Journal of Agricultural Economics*, 26(1):1–13, April 1982. ISSN 0004-9395, 1467-8489. doi: 10.1111/j.1467-8489.1982.tb00404.x. URL http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1111/j.1467-8489.1982.tb00404.x/abstract.

[ ]Richard Schmalensee. Appropriate government policy toward commercialization of new energy supply technologies. *Energy Journal*, 1(2):1–40, 1980. ISSN 0195-6574. URL http://www.jstor.org.proxy.bc.edu/stable/41321461.