

Privacy Externalities and “Opt Out”: Theory and Evidence from Do Not Call

Kai-Lung Hui[†], I.P.L. Png^{*}, and Khim-Yong Goh^{*}

August 2011

Abstract

Concerned about privacy, consumers have opted out from direct marketing by the millions. Legislators have proposed laws to enable consumers to opt out from monitoring of browsing activity. However, there has been relatively little analysis of how opt out facilities actually work in practice. Here, we make two contributions. First, we develop an analytical model of opt out in the context of consumer privacy that explicitly accounts for strategic behavior by sellers and consumers. We show that opt out increases with the consumer population if the effect of consumer opt out on vendors outweighs that on consumers. Second, we take the analysis to data from the U.S. Do Not Call registry, using two empirical strategies to test the externality. One is a cross-section analysis with the consumer population instrumented by land area. The other exploits the exogenous timing of the registry for a difference-in-differences test. Supported by multiple robustness checks and falsification tests, we conclude that opt out facilities do give rise to externalities between consumers and vendors. Further, the externalities result in a diseconomy of scale in direct marketing.

[†] Hong Kong University of Science & Technology; ^{*}National University of Singapore. We thank Hal Varian for pointing us to the federal Do Not Call registry. We gratefully acknowledge financial support from the NUS Academic Research Fund (Grants R253-000-030-112 and R253-000-034-112,) and the Hong Kong SAR Central Policy Unit and Research Grants Council (HKUST 1001-PPR-3).

1 Introduction

The explosive growth of social media and the consequent trail of Web browsing presents both opportunities and risks – opportunities for businesses to refine their marketing and risks for consumers of harm to privacy. The risks have prompted consumers to press for government action. Caught in a bind between growth of online business and protection of privacy, regulators and legislators have sought a middle way. On the one hand, they wish to avoid prescriptive regulation that would curb legitimate online businesses. On the other hand, they must respond to consumer concerns for privacy.

One seemingly promising middle way is to provide consumers with facilities to “opt out”. In principle, opt out would ensure economic efficiency. Consumers who are most harmed would opt out. Consumers who benefit relatively more from online businesses would remain available for marketing. To the extent that consumer distaste for marketing aligns with their responsiveness to marketing offers, the opt out facility actually helps businesses. It reduces the waste from marketing to consumers who are anyway not interested in the offers (Johnson 2003; Rotfeld 2004).

Opt out mechanisms have been implemented in the contexts of direct mail and tele-marketing, and proven extremely popular with the public in the United States, Australia, United Kingdom, and elsewhere. With increasing concern about harms to privacy from tracking over the Internet, legislators have pushed laws to provide consumers with facilities to opt out of their browsing activity being monitored. So-called “do not track” bills have been introduced in the U.S. House of Representatives by Ms Jackie Speier and in the California Senate by Mr Alan Lowenthal.

Although opt out mechanisms are widely supported by regulators and legislators as a balanced way to address privacy, there has been relatively little analysis of how they actually work in practice. Analytical research into avoidance of marketing in general has suggested that consumer actions to avoid marketing may generate externalities for each other (Hann et al. 2008; Johnston 2008; Armstrong et al. 2009). However, the scant empirical evidence has focused on the influence of demographics and publicity on opt-out behavior, with no attention to the possible externalities (Varian et al. 2004, 2005; Goh et al. 2011).

Here, we make two contributions. First, we develop an analytical model of opt out in the context of consumer privacy that explicitly accounts for strategic behavior by sellers and consumers. By contrast with previous analytical research, which studied marketing avoidance in general, our analysis focuses specifically on opt out. We show that opt out

generates externalities among consumers: When one consumer opts out, she shrinks the pool available for marketing, and so, increases the expected harm for others. Further, sellers re-direct solicitations as consumers opt out. We derive a particular empirical implication – that opt out increases with the consumer population if the effect of consumer opt out on vendors outweighs that on consumers.¹

Second, we take the empirical implications of our analysis to data from the U.S. Do Not Call (DNC) registry. We use two empirical strategies to test the externality. One approach uses a cross-section analysis, with the consumer population instrumented by land area. We estimate the elasticity of registrations with respect to number of households to be 0.26. The other approach exploits the exogenous timing of the registry to design a difference-in-differences test. We compare the pattern of DNC registrations before and after September 2, 2003, and find that DNC registration increased with the consumer population from September 2. The findings from the two approaches are buttressed by multiple robustness checks and falsification tests.

Our empirical findings of externalities among consumers and sellers provide insight and guidance to both managerial practice and public policy. Managers need to appreciate how the yield from solicitations varies with consumer and seller adjustment to an opt out facility. The externalities give rise to a diseconomy of scale, making larger markets relatively less attractive. Considering the complexity of strategic interaction, policy makers might give more consideration to policies that more directly address the harm caused by solicitations, such as taxes (Hann et al. 2008).

2 Model

Consider a market with sellers $m = 1, \dots, M$, and potential consumers, $i = 1, \dots, N$. Each seller m makes S_m solicitations to sell some item at a fixed price, p . Each potential consumer i can buy item m only if solicited by seller m and values the item at v_{im} . So, her net benefit from the item is $b_{im} = \max\{v_{im} - p, 0\}$.

Consumers are risk neutral and suffer harm, w , from each solicitation received. A consumer can opt out of all direct marketing solicitation by incurring a cost, e , to register

¹Our analytical modeling applies the methods of supermodularity to derive the comparative statics and empirical implications (Van Zandt 2002). These methods are especially useful in contexts where, owing to economies of scale, the convexity typically assumed in classical microeconomics may not hold. The methods are useful in a broad range of management research (Milgrom and Shannon 1994).

with a “do not contact” (DNC) list. We assume that the cost of opting out is identical for all consumers. Once the consumer has opted out, she will not receive any solicitations and will avoid any harm.

Given the list of consumers who have *not* opted out, the cost to seller m of making S_m solicitations is $C(S_m)$. The seller’s cost includes the costs of collecting consumers’ contact information, checking the contacts against any DNC registry, and making the solicitations. We assume that each seller programs its solicitations such that it does not contact any consumer twice. For simplicity, the cost of producing the item is zero.

The timeline is as follows. First, consumers decide whether to opt out (i.e., register with the DNC list), and, simultaneously, sellers make solicitations.² Consumers i who have not opted out will buy item m provided that they receive a solicitation from seller m and their valuations for the item exceed the price, i.e., $v_{im} \geq p$.³

Let n be the number of consumers who choose *not* to opt out and so are available for solicitation. Consider consumer i . She must choose between: (i) opting out by registering for DNC, avoiding all solicitations, and getting net benefit, $-e$, or (ii) not opting out and getting net benefit in (1) below. If she does not opt out, she will buy item m if she receives a solicitation from seller m and $v_{im} \geq p$. Her expected net benefit from receiving solicitations is the sum of her expected net benefits from all sellers, $m = 1, \dots, M$,⁴

$$B_i(S_1, \dots, S_M) \equiv \sum_{m=1}^M \frac{S_m}{n} [b_{im} - w]. \quad (1)$$

Denote the consumer’s decision to opt out by r_i . Consumer i would choose to opt out, $r_i = 1$, if $B_i \leq -e$, and not opt out, $r_i = 0$, otherwise. Thus, her choice of whether to opt out,

$$r_i^* = \arg \max_{r_i \in \{0,1\}} \{[1 - r_i]B_i(S_1, \dots, S_M) - r_i e\}, \quad (2)$$

²In practice, sellers must check their contact lists against the DNC registry before making the solicitations, whereas consumers would make the opt out decision by anticipating the number of solicitations that the sellers will make to the pool of unregistered consumers. Analytically, such a decision structure with consumer and seller expectations is equivalent to a setting in which the sellers and consumers make simultaneous decisions.

³In related work, Hann et al. (2008) model consumer efforts to avoid marketing and privacy externalities. In a model with two types of consumer and essentially only one item for sale, they study social welfare and public policy. By contrast, here, we allow multiple consumer types and sale of multiple items, and focus on empirical implications of privacy externalities arising from opt out.

⁴Recall that a seller would contact a consumer at most once. Hence, the consumer would incur harm, w , at most once from each seller.

where $B_i(S_1, \dots, S_M)$ is defined in (1). Evidently, the opt out decisions of others exert an externality directly through the number of consumers who choose *not* to opt out, n , and indirectly through seller solicitations, S_m .

The N consumers can be ordered by their expected net benefit, $B_i(S_1, \dots, S_M)$. Accordingly, redefine the index, $i = 1, \dots, N$, in increasing order of the expected net benefit. Let λ be the marginal consumer, such that consumers who derive relatively lower benefit, $i \leq \lambda$, opt out, whereas those who derive larger benefit, $i > \lambda$, do not opt out. Then, equation (2) characterizes the consumers' reaction functions, which indirectly characterizes the marginal consumer, λ . Further, by definition,

$$n = N - \lambda. \quad (3)$$

The following result shows that the extent of opt out, λ , increases with seller solicitations, S_m .⁵ Intuitively, for the infra-marginal consumer, $\lambda + 1$, her expected benefit from solicitations is relatively low compared to the harm. So, an increase in seller solicitations would increase her expected harm relatively more than her expected benefit, and so, tip her toward opting out. Similarly, with more consumers available for solicitation, the harm incurred on each consumer decreases. An increase in the harm caused by solicitations would cause more consumers to opt out.

Proposition 1 *For the marginal consumer, λ , if $b_{\lambda m} < w$, $m = 1, \dots, M$, then consumer opt out*

- (i) *increases with seller solicitations, S_m ,*
- (ii) *decreases with the number of consumers who choose not to opt out, n , and*
- (iii) *increases with the harm caused by solicitations, w .*

Note that, if $S_m = 0$, $m = 1, \dots, M$, the consumer's expected net benefit from direct marketing is zero, and so, $\lambda = 0$. As S_m increases, some low benefit consumers i would opt out because $B_i < -e$. If the population, N , is large and consumers' net benefits from the marketed items, b_{im} , $i = 1, \dots, N$, are sufficiently dispersed, then there always exists a marginal consumer, $\lambda < N$, so that, $B_\lambda \leq -e$, and $B_{\lambda+1} > -e$. Proposition 1 states that the *rank* of such a marginal consumer, out of the N consumers, increases in S_m and w , and decreases in n .⁶

⁵Please refer to the Appendix for the proofs of all results.

⁶Proposition 1 uses the condition, $b_{\lambda m} < w$, for all m . As long as $B_\lambda \leq -e$, parts (ii) and (iii) would continue to hold if some $b_{\lambda m} \geq w$. Part (i) may not hold for some m' if $b_{\lambda m'} \geq w$. However, as long as $B_\lambda \leq -e$, λ would increase with a concomitant increase in all S_m , $m = 1, \dots, M$, which generates the same empirical implication as stated in part (i) of Proposition 1. For tractability, we use the stronger but simpler condition, $b_{\lambda m} < w$, in Proposition 1.

The broken curve in Figure 1 graphs the marginal consumer, λ , against the solicitations of one seller, m . For simplicity, we treat λ and S_m as continuous variables.⁷

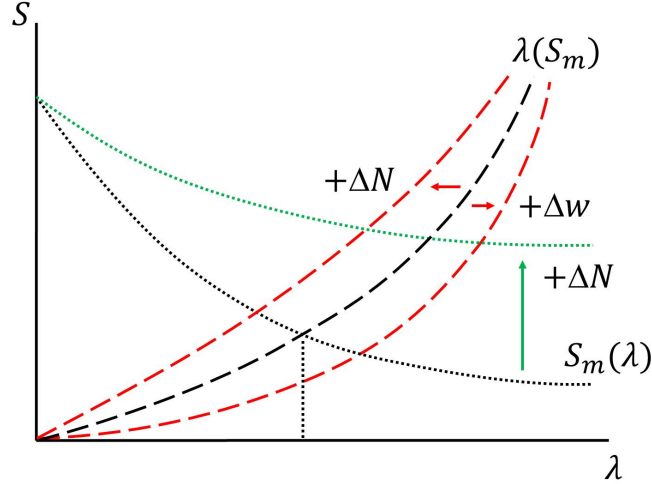


Figure 1: Consumer Registration and Seller Solicitations

We next derive the seller's reaction function, $S_m(\lambda)$. Seller m 's expected revenue from consumers $i = \lambda + 1, \dots, N$ is,⁸

$$R(S_m) = \frac{S_m}{n} \sum_{i=\lambda+1}^N pq_{im}, \quad (4)$$

where $q_{im} = 1$ if $v_{im} \geq p$, and $q_{im} = 0$ otherwise. Taking into account the cost of solicitation, seller m 's expected profit is

$$\Pi_m(S_m) = \frac{S_m}{n} \sum_{i=\lambda+1}^N pq_{im} - C(S_m). \quad (5)$$

Hence, its choice of solicitation,

$$S_m^* = \arg \max_{S_m} \left\{ \frac{S_m}{n} \sum_{i=\lambda+1}^N pq_{im} - C(S_m) \right\}. \quad (6)$$

If fewer consumers choose to opt out (i.e., as n increases), the seller's expected revenue would be higher, and so it would send more solicitations.⁹ Further, sellers ignore the harm that their solicitations impose on consumers, and so, solicitations are independent of the harm. Our next proposition follows.

⁷Figure 1 should have $M + 1$ dimensions. For ease of presentation, and, without loss of generality, we draw the reaction function for only one seller, m .

⁸By definition of λ , the pool of n consumers available for marketing is those indexed $i = \lambda + 1, \dots, N$.

⁹Hann et al. (2008) show that if some low-benefit consumers do not consider a marketed item, then it is possible that sellers *reduce* solicitations as such low-benefit consumers choose not to opt out.

Proposition 2 *Seller solicitations, S_m ,*

- (i) *weakly increase with the number of consumers who choose not to opt out, n , if $v_{\lambda m} \geq p$,
and*
- (ii) *are independent of the harm caused by solicitations, w .*

Given an opt out facility, Propositions 1 and 2 characterize how changes in n and w affect the reaction functions of consumers and sellers. Given the population, N , by (3), $n = N - \lambda$. Hence, Propositions 1 and 2 *indirectly* characterize how changes in the population size, N , and the opt out rate, λ , affect consumers and sellers. Figure 1 illustrates the equilibrium consumer opt out and seller solicitation. The downward-sloping dotted curve represents the seller reaction function, i.e., the seller's solicitation, S_m , as a function of the marginal consumer, λ .¹⁰

We can use Figure 1 to motivate the *net* impact of changes in the consumer population, N , and the harm caused by solicitations, w . In particular, an increase in the consumer population would directly increase the number of consumers who do not opt out and are available for solicitation, n . This has two different effects:

- (a) *Dispersion effect*: The average harm caused by seller solicitations on each consumer decreases. Fewer consumers would opt out (Proposition 1, part (ii)).
- (b) *Sales effect*: There are more potential consumers in the market, and so sellers increase solicitations (Proposition 2, part(i)), so, raising consumer opt out.

Referring to Figure 1, as the shifts in the curves show, the net effect could be to reduce or increase opt out, depending on whether the effect on consumers (dispersion effect) or the effect on sellers (sales effect) is stronger. Accordingly, the impact of the consumer population size on consumer opt out is an empirical question.

Hypothesis 1 *Consumer opt out increases with the consumer population if the sales effect outweighs the dispersion effect.*

By contrast, the effects of an increase in the harm caused by solicitations is unambiguous. Consumer opt out would increase (Proposition 1), but there would be no effect on seller solicitation (Proposition 2). Hence, the result would be more opt out.

Hypothesis 2 *Consumer opt out increases with the harm caused by solicitations.*

¹⁰Part (i), Proposition 2 uses the condition, $v_{\lambda m} \geq p$. If $v_{\lambda m} < p$, then referring to Figure 1, the S_m curve would slope upward instead of downward in λ (cf. Hann et al. 2008, Proposition 1), but it would not affect the strategic analysis and the subsequent empirical results.

3 Context and Data

With limited exceptions, federal law prohibits unsolicited telemarketing calls to telephone numbers on the U.S. DNC registry.¹¹ The Federal Trade Commission (FTC) administers the U.S. DNC registry, which was opened to consumers on June 27, 2003. Registrations before September 1 were effective from October 1. Specifically, telemarketers were allowed to access the registry from September 2 and required to stop calls to registered numbers no later than October 1. DNC registrations from September 1 onward were effective after a 90-day processing period.

We chose to conduct the analysis at the county level for several reasons. Telemarketing is governed by state-level laws and regulations, and so it is important that the unit of analysis lie within state boundaries. While counties fit within state boundaries, telephone area codes may not. Further, we needed demographic information, which were available at the county level from the U.S. Census.

The FTC provided us with redacted telephone numbers on the DNC registry for each area code and exchange, for example, (617) 363-xxxx, by date of registration. On average, each U.S. county spanned 47 exchanges, but there was substantial variation in the number of exchanges per county. We allocated the DNC registrations to counties using the *North American Local Exchange NPA-NXX Database*. We matched the DNC registrations with county-level population characteristics from the 2000 U.S. Census.¹²

Prior to the federal DNC registry, 27 states had already established a similar “do not call” facility (Varian et al. 2004). Of these, 16 states eventually merged their lists with the federal registry. The presence of a state list would reduce the consumer’s benefit from the federal DNC registry.

Figure 2 presents the geographical distribution of DNC registration rates. Evidently,

¹¹The exceptions are: calls for political campaigning and survey research, by nonprofit and charitable organizations, and by businesses with a recent commercial relationship with the consumer. The federal DNC Registry applies to interstate and intrastate telemarketing calls and accepts registrations from fixed-line and mobile but not business telephone numbers. Telemarketers are required to remove numbers on the DNC Registry from their call lists no less frequently than every 31 days. Violators are subject to fines of US\$11,000 per offence.

¹²On reviewing the data, we found that there were less than one DNC registration in Kenedy, TX (FIPS 48261) and Loving, TX (FIPS 48301), and Williamsburg, VA (FIPS 51830) had an average registration of 4.73 per household. These numbers were exceptional, and so we excluded these three counties from subsequent analyses.

there was considerable variation in DNC registration by state, some of which might be due to the presence of a state DNC registry.

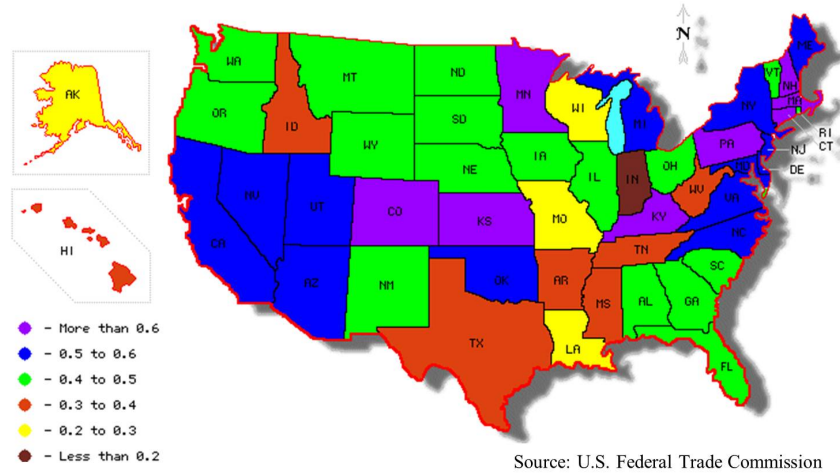


Figure 2: DNC registration rates (per household), June 2004

Figure 3 presents the daily DNC registrations in states without state lists. There were several waves of registrations in the first few months after June 27, 2003.

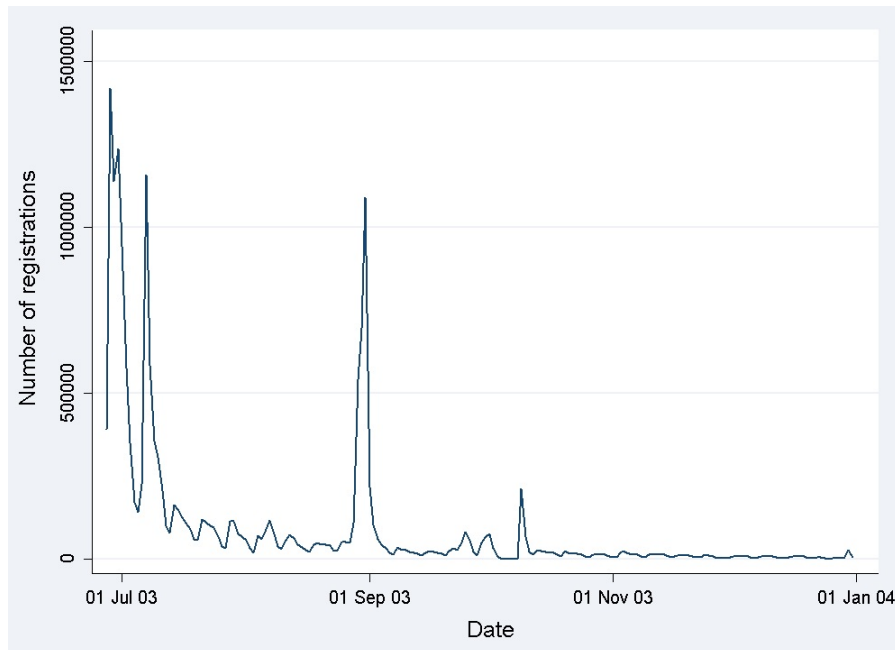


Figure 3: Number of DNC registrations over time

4 Empirical Strategy

We applied two different identification strategies to estimate the impacts of the number of consumers and the harm caused by solicitations on registrations. The first analyzed cross-sectional variation in registrations using an instrumental variable. The second exploited an exogenous difference in timing between the opening of the DNC registry to consumers and to telemarketers.

The first strategy related the cumulative registration rate to a set of variables representing the consumer population and the harm caused by solicitations. Specifically, we estimated the following cross-sectional equation,

$$R_k = \alpha_1 N_k + \alpha_2 X_k + \epsilon_k, \quad (7)$$

where R_k is the overall DNC registration rate, i.e., the cumulative number of registrations divided by the number of households, λ/N , in county k , N_k is the number of households in county k , X_k are county characteristics that represent the harm caused by solicitations and control for other county-specific effects, and ϵ_k is a random county-level error. By Hypothesis 1, we expect $\alpha_1 > 0$ if the sales effect outweighs the dispersion effect, and $\alpha_1 < 0$ otherwise.

The challenge in identifying α_1 , however, is that R_k and N_k may be correlated for other reasons. There may be economies of scale in solicitations, and so, markets with larger populations attract more telemarketers and the greater volume of solicitations induces more consumers to opt out. Other contextual factors might also explain the relation between the population and registrations (Manski 1993). For example, consumers are more likely to communicate with like-minded people in a larger population (Alesina and La Ferrara 2000; Marmaros and Sacerdote 2006).

Following Angrist and Krueger (2001), we sought an instrument to provide exogenous variation in the problem variable, consumer population, N_k . Land area is exogenously correlated with consumer population. A larger land area can naturally accommodate more consumers. However, land area is hardly related to telemarketing practices or consumer social interactions. The essential idea is that nature would randomly assign consumers to different regions according to space. Therefore, land area would likely correlate with the endogenous regressor, N_k , but not the outcome variable, R_k .

To lend confidence to our choice of land area as the instrument, Figure 4 plots the number of households and DNC registration rates at the end of June 2004 against county

land area.¹³ Evidently, counties with larger land area tend to have more households, but they do not seem to exhibit higher DNC registration rates.

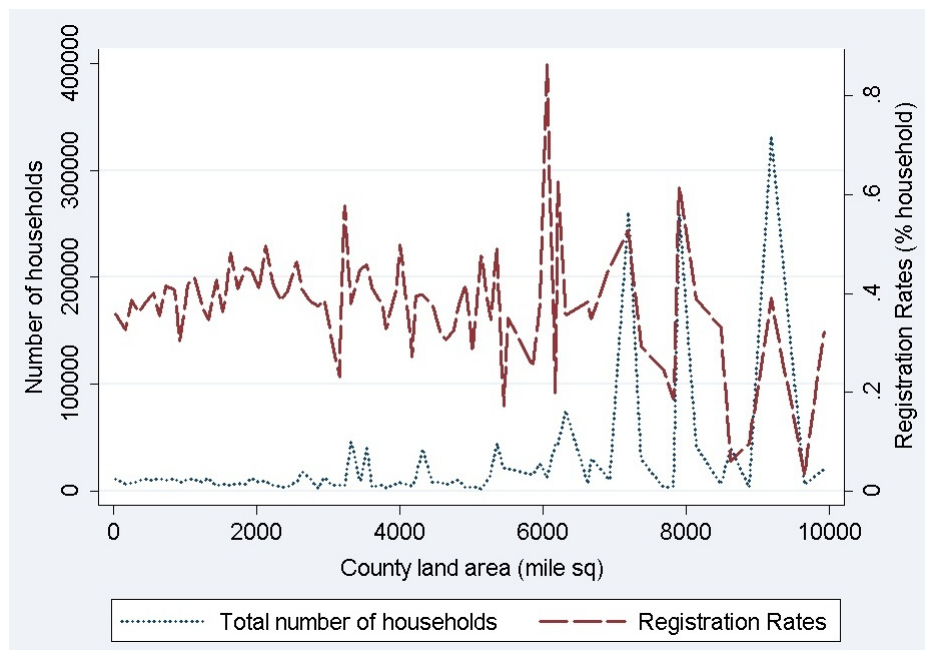


Figure 4: DNC registration rates and number of households, by county land area

The cross-sectional estimation of (7) allowed us to identify the equilibrium relationship between the number of consumers and DNC registration rates. This approach is subject to two limitations. First, even instrumenting for the consumer population, we still may not completely account for endogenous factors. Second, due to its static nature, we could not study the dynamics of consumer registration and seller solicitation.

Our second identification strategy exploited the exogenous difference in timing between the opening of the DNC registry to consumers and to telemarketers. By design, the difference-in-differences strategy ruled out endogenous factors. Further, it allowed us to study the dynamics of consumer registration, and, implicitly, seller solicitation.

The FTC opened the DNC registry to telemarketers on September 2, 2003, and required them to stop calls to numbers registered before September 1 no later than October 1. So, before September 1, consumers registered for DNC based on their *expectation* of consumer and seller behavior. The externalities, if any, would have materialized only after

¹³We divided the counties into 100 bands of equal width according to land area. Figure 4 plots the median number of households and DNC registration rate in each band of counties. As most counties were small in land area, a scatter plot would be overly concentrated at the lower tail, and so would not reveal the patterns clearly.

September 2, as telemarketers refined their calling lists to remove telephone numbers registered with the DNC list. In particular, the *sales* effect would have materialized only when telemarketers shifted their calling to numbers which were not registered.¹⁴

Figure 5 shows the timing of the implementation of the DNC registry. Telemarketers could access the registry and refine their calling lists from September 2 (shaded region). It is clear that any externalities arising from the *sales* effect could materialize only from September 2. This motivates a difference-in-differences estimation:

$$r_{kt} = \beta_1 N_k \cdot \text{SEPT}_t + \beta_2 X_{kt} + \beta_3 \psi_k + \beta_4 \tau_t + e_{kt}, \quad (8)$$

where r_{kt} is the DNC registration rate (registrations divided by number of households) in county k on day t , the indicator, $\text{SEPT}_t = 1$ for any day on or after September 2 and $\text{SEPT}_t = 0$ otherwise, X_{kt} are other factors, including variables representing the harm caused by solicitations, that might possibly affect DNC registrations, ψ_k are county-level fixed effects, τ_t are day fixed effects, and e_{kt} is an idiosyncratic error.¹⁵

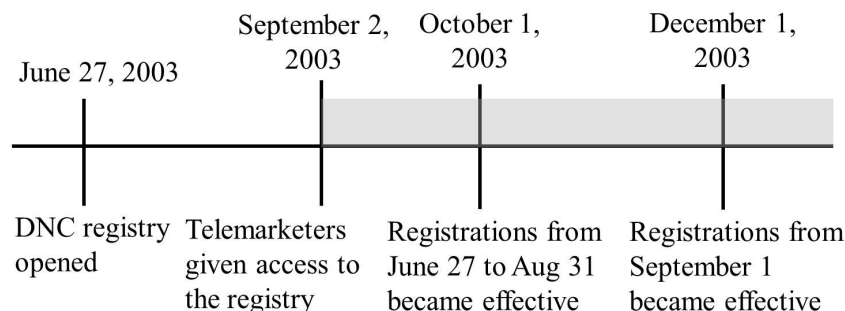


Figure 5: Timing of DNC implementation

With daily registration data and the exogenous September 2 opening of the registry to telemarketers, we can unambiguously identify the impact of seller solicitations. In particular, if $\beta_1 > 0$, it indicates that consumers in larger markets were more likely to register with DNC after September 2. From our review of the relevant literature (FTC 2003; Johnson 2003; Rotfeld 2004; Varian et al. 2004 and 2005), telemarketer access to the DNC registry was the only relevant event around September 2. Hence, we interpret $\beta_1 > 0$ as the sales

¹⁴Within the month of September, 13,000 telemarketing organizations subscribed to the DNC Registry. Among those, over 400 paid for access to the entire registry (FTC 2003). DNC registrations from September 1 onward were effective only after a 90-day processing period.

¹⁵The effect of SEPT_t itself is captured by the day fixed effects, τ_t , and so cannot be separately identified. Also, for ease of interpretation, we specified all variables that were interacted with a time indicator, such as September 2 and after, as their deviation from the mean (Echambadi and Hess 2007).

effect (increased telemarketing calls from sellers to unregistered consumers) outweighing the dispersion effect (reduced harm due to larger consumer population). By contrast, if $\beta_1 < 0$, then the realized sales effect was outweighed by the dispersion effect.

We specified equation (8) as a daily panel, with the cumulative registration rate up to day $t - 1$ as a control variable in X_{kt} . This would account for the diffusion of DNC registrations over time (Mahajan et al. 1990; Van den Bulte and Stremersch 2004). As more consumers register with the DNC list, the number available to register falls, and so, the registration rate would fall. The cumulative registration would also control for unobserved consumer heterogeneities across counties and time, such as social learning, peer influence via communication, homophily, or herding.

Figure 3 clearly shows that DNC registrations peaked at multiple times, which is difficult to capture with a simple model (Bonfrer and Dreze 2009). Hence, to identify the effect of the September 2 shock, in estimating (8), we limited the panel to 14 days, from 7 days before September 2 to 7 days after. We specified all continuous variables in logarithms as the variables have widely differing scales and variations (Wooldridge 2006).

5 Results

5.1 Cross Sectional Analysis

Table 1 presents summary statistics of registrations and the covariates.¹⁶ We used household size and unemployment rate to represent the harm caused by solicitations. People in larger households would be less likely to pick up telemarketing calls, and so, they should suffer less harm from telemarketing. By contrast, unemployed people have a lower cost of time and should suffer less harm from telemarketing. By Hypothesis 2, household size and unemployment rate should be *negatively* correlated with consumer DNC registrations.¹⁷

Table 2, column (1), reports the ordinary least squares (OLS) estimates of (7), which included a full set of state fixed effects and were estimated with robust standard errors

¹⁶Note that the DNC registration rate could exceed one because some households may own multiple telephone lines. The rate would be bounded by one if we set N as the number of telephone lines. However, we could not find the number of telephone lines by county.

¹⁷The cross-sectional analysis applies to the static equilibrium between consumers and vendors. Accordingly, we included all counties, including those in states with state-level DNC registries. The static equilibrium would represent all externalities, whether generated by federal- or state-level registries.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
DNC registration rate (λ/N)	0.396	0.192	0.002	1.624	3125
No. of households ('000)	33.699	104.454	0.185	3133.774	3125
County land area (mile ²)	923.687	1085.623	1.986	9953.182	3078
Household size	2.630	0.239	2.073	5.127	3125
Household income (\$'000)	35.354	8.866	12.692	82.929	3125
Unemployed (%)	3.438	1.522	0	32.863	3125
Commute time (mins)	23.425	5.653	6.275	48.667	3125
Retail density (stores per square mile)	0.398	1.610	0	61.438	3078

clustered by state (Wooldridge 2006). The state fixed effects help control for state-specific telemarketing regulations and other unobserved state characteristics. The coefficient of number of households was positive and significant ($\alpha_1 = 0.096$, $p < 0.01$). Apparently, DNC registration *increased* with the consumer population, which is consistent with the sales effect outweighing the dispersion effect in Hypothesis 1. Further, both household size and unemployment rate were negatively related to registrations, which is consistent with Hypothesis 2.

As discussed above, the OLS estimate of α_1 may be biased by unobserved economies of scale in telemarketing or endogenous social processes. To overcome this problem, we instrumented population by the land area of the county. We first checked the validity of land area as an instrument. In the first-stage regression, land area was strongly predictive of the number of households (Cragg-Donald F statistic = 21.40, $p < 0.01$). This indicates that land area passes the weak instrument test (Staiger and Stock 1997; Stock and Yogo 2005), and so meets the *relevance* condition for instrumental variable estimation.¹⁸ The *exclusion* restriction for instrumental variable estimation is generally untestable (van den Berg 2006), but the plots in Figure 4 show that there was little correlation between land area and DNC registrations. Hence, we concluded that land area was a valid instrument for population.

Table 2, column (2), presents the 2SLS (two-stage least squares) estimates. The coefficient of the number of households was positive and significant, and was almost three times

¹⁸Specifically, the Cragg-Donald F statistic exceeded the critical value for the 10% maximal size in the Stock and Yogo (2005, Table 5.2) weak instrument identification test. Further, the Anderson-Rubin Wald statistic, $\chi^2 = 24.25$ ($p < 0.01$), indicating that the estimated coefficient of number of households was robust to weak instruments. We also conducted Wooldridge's (1995) score test and concluded that the number of households was indeed endogenous ($F = 11.61$, $p < 0.01$).

Table 2: Cross-Sectional Regressions

VARIABLES	(1) OLS	(2) 2SLS	(3) 2SLS: MSA counties	(4) 2SLS: Non-MSA counties
No. of households	0.096*** (0.014)	0.260*** (0.062)	0.593* (0.336)	0.180*** (0.029)
Household size	-1.149*** (0.154)	-1.046*** (0.201)	0.874 (1.469)	-1.118*** (0.153)
Household income	0.372*** (0.078)	-0.228 (0.236)	-1.451 (1.089)	0.477*** (0.112)
Unemployed	-3.353*** (0.998)	-8.416*** (2.258)	-26.138 (19.024)	-4.827*** (1.097)
Commute time	-0.057 (0.070)	-0.006 (0.068)	0.693** (0.292)	-0.033 (0.070)
Retail density	-0.064 (0.061)	-0.136*** (0.045)	0.126 (0.169)	-0.165*** (0.051)
Constant	-4.551*** (0.817)	-0.017 (1.921)	5.725 (6.593)	-6.447*** (1.165)
Observations	3,078	3,078	841	2,237
R-squared	0.483			

Notes. The dependent variable is log cumulative registration rate. All specifications included state fixed effects. Robust standard errors clustered by state in parentheses. Column (1): OLS estimates. Column (2): 2SLS estimates. Column (3): 2SLS estimates for MSA counties. Column (4): 2SLS estimates for non-MSA counties.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

the OLS estimate in column (1). The elasticity of registrations with respect to the number of households was 0.26 (± 0.062). Apparently, market size, as measured by the county population, did significantly affect DNC registrations. By Hypothesis 1, the result was consistent with the sales effect outweighing the dispersion effect. Both household size and unemployment rate were negative and significant, which was consistent with Hypothesis 2.

To further check our identification strategy, we conducted an additional falsification test. The use of county land area as an instrument rests on the assumption that nature would randomly locate more consumers in a larger region. This, however, may not be true for metropolitan areas which attract more consumers regardless of land area. If land area is also related to population and DNC registrations in metropolitan areas, then county land

area would not be a suitable instrument.

Table 2, columns (3) and (4), report 2SLS estimates for counties in metropolitan statistical areas (MSA) and other areas. The coefficient of number of households was only significant among non-MSA counties. It was insignificant for counties in MSAs. Indeed, among MSA counties, in the first stage regression of number of households on land area, the coefficient of land area was 0.219, with $p > 0.05$. This suggests that the correlation between county land area and number of households in our full sample was not due to factors other than space.

5.2 Difference-in-Differences

Here, we exploited the exogenous timing of the federal DNC registry, being opened to consumers from June 27, 2003, and to telemarketers only from September 2, 2003, as a natural experiment. Specifically, we estimated (8) using daily DNC registration data for 14 days around September 2 (± 7 days). Since consumers in states with a pre-existing state registry might have already experienced the externalities and so would be less sensitive to any change in seller solicitations after September 2, we excluded counties with state-level DNC registries from this set of analyses.

Table 3 reports the OLS estimation results, including a full set of county and day fixed effects, with robust standard errors clustered by county. As a baseline, column (1) reports an estimate including only the time-varying control and demographic variables.¹⁹ The most important control variable was the cumulative DNC registration rate in the county up to day $t - 1$. The cumulative DNC registration rate controlled for the diffusion of registration over time, as well as unobserved heterogeneity. As expected, the coefficient of the cumulative registration rate was negative and significant.

The other time-varying control variable was the circulation-weighted number of newspaper reports of DNC per household in the county by day (Goh et al. 2011). As intuitively expected, the coefficient of news reports was positive, albeit insignificant.²⁰ All non-time-varying and non-county-varying factors that would possibly influence DNC registrations

¹⁹For all non-time varying variables (household size, household income, etc.), we included their interaction with the indicator of September 2 and after as control variables. For brevity, we do not report these control variables.

²⁰This coefficient probably under-estimates the impact of news reports. News reports might affect registration over time. Not every consumer reacts immediately. Indeed, in a weekly specification, news reports were significantly associated with DNC registrations (Goh et al. 2011).

Table 3: Daily Panel Regressions

VARIABLES	(1) Baseline	(2) Extern- alities	(3) Robustness: Social Interaction	(4) Robustness: Telemarketing Sales	(5) Robustness: Race Hetero	(6) Robustness: Ling Isol
Cumulative Registrations [t-1]	-0.144*** (0.004)	-0.065*** (0.008)	-0.064*** (0.008)	-0.064*** (0.008)	-0.064*** (0.008)	-0.061*** (0.008)
Cum reg \times Sep 2		-0.025*** (0.002)	-0.025*** (0.002)	-0.025*** (0.002)	-0.024*** (0.002)	-0.025*** (0.002)
No. of households \times Sep 2		0.267*** (0.067)	0.242*** (0.068)	0.302*** (0.067)	0.232*** (0.073)	0.311*** (0.072)
News reports	0.026 (0.019)	0.014 (0.017)	0.015 (0.017)	0.015 (0.017)	0.014 (0.017)	0.018 (0.018)
Social Interaction \times Sep 2			-0.339*** (0.115)			
Telemarketing Sales \times Sep 2				-0.241*** (0.059)		
Race Heterogeneity \times Sep 2					0.902** (0.391)	
Ling Isolated \times Sep 2						-0.021 (0.048)
Constant	42.108*** (1.146)	19.803*** (2.344)	19.957*** (2.337)	19.507*** (2.345)	19.608*** (2.315)	18.689*** (2.276)
Observations	16,506	16,506	16,506	16,506	16,506	15,848
R-squared	0.806	0.816	0.816	0.816	0.816	0.819
Number of fips	1,179	1,179	1,179	1,179	1,179	1,132

Notes. The dependent variable is log registration rate. For ease of presenting the estimates, we rescaled log registration rate, log cumulative registration rate, and log news reports per household by multiplying them by 1000. All specifications included the interactions between September 2 and household size, household income, unemployment, commute time, and retail density as control variables (not reported), and county and day fixed effects. Robust standard errors clustered by county in parentheses. Column (1): Baseline OLS estimates with cumulative registrations, news reports, and demographics \times September 2. Column (2): Including externalities \times September 2. Column (3): Including social interaction \times September 2. Column (4): Including telemarketing sales \times September 2. Column (5): Including race heterogeneity \times September 2. Column (6): Including percent linguistically isolated \times September 2.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

would have been absorbed by the county and day fixed effects, and so could not be separately estimated.

Table 3, column (2), reports our main estimate including the identification of the externalities. To check the validity of our specification, Figure 6 plots the overall DNC registration rates and the estimated day fixed effects over time. Evidently, the day fixed effects captured the time profile of registrations, so, our estimates focused on the *difference* between registrations and the average for the day. Hence, any observed effect of the number

of households interacted with time was not due to unobserved day-specific shocks.

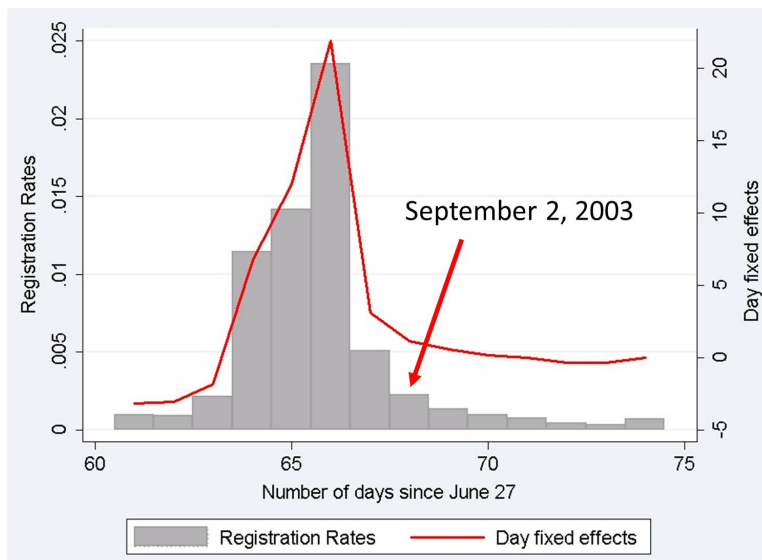


Figure 6: DNC registrations and day fixed effects

Indeed, as shown in Table 3, column (2), the coefficient of the number of households interacted with the time indicator of September 2 and after was 0.267 (± 0.067), which was positive and significant. Apparently, on and after September 2, for a county with a population that was 1% (around 340 households) larger, the daily registration rate would be higher by 0.267%. This is strong evidence of the externality due to seller solicitations. In counties with larger populations, vendors sent relatively more solicitations to consumers, causing more consumers to register for DNC.²¹ Referring to Hypothesis 1, this result suggests that the sales effect due to increased seller solicitations outweighed the dispersion effect due to a larger consumer population.

The coefficient of the cumulative registration rate interacted with the indicator of September 2 and after was negative and significant. Apparently, the time-profile of registrations tended to fall but at a diminishing rate over time. Hence, cumulative registrations had a lower impact on daily registrations after September 2.²²

²¹It is instructive to observe from Figure 6 that, shortly before September 2, there was a sharp peak of DNC registration. Supposing the day fixed effects did not sufficiently control for the peak, the coefficient of the number of households interacted with the indicator of September 2 and after would be biased *downward*. Hence, our observation that the number of households interacted with the indicator of September 2 and after had a significant and *positive* coefficient suggests that the day fixed effects had effectively controlled for time-specific events affecting DNC registrations.

²²This interpretation is supported by a set of falsification tests, as reported in Table 4, using August 2 and October 2 as placebo cut-off dates; see the discussion below.

To buttress the finding of an externality, we conducted various robustness tests, as reported in Table 3, columns (3)-(6). The first was to test the alternative explanation that social influence caused the increased registration after September 2 in more highly populated counties (Bertrand et al. 2000; Tucker 2008). In particular, it may be easier for consumers in a larger market to communicate with each other about changes in telemarketing solicitations before and after September 2.

We obtained data on social interaction from the Social Capital Community Benchmark Survey.²³ We multiplied the social interaction variable by the indicator of September 2 and after, and entered the product as an additional covariate in estimating (8). As reported in Table 3, column (3), our main finding was robust to this additional covariate. The coefficient of the number of households interacted with the indicator of September 2 and after was positive and significant, albeit somewhat smaller than the main estimate. Interestingly, the coefficient of social interaction interacted with the indicator of September 2 and after was negative and significant, possibly because more socially active consumers were less affected by telemarketing calls at home.

Our next two robustness tests addressed the possibility that the county population might have affected DNC registrations through telemarketing practices. The county population could have represented some latent effects that also affected the extent to which telemarketers send solicitations (e.g., consumer preferences for direct marketing), in which case, the population may be correlated with telemarketing expenses. We obtained the state-level consumer telemarketing sales in 2003 from the Direct Marketing Association (2006). Table 3, column (4), reports an estimate including telemarketing sales multiplied by the indicator of September 2 and after as an additional covariate. The coefficient of the number of households interacted with the time indicator of September 2 and after was positive and significant, and larger than the main estimate. The coefficient of telemarketing sales multiplied by the indicator of September 2 and after was negative and significant, suggesting that, on and after September 2, DNC registrations fell relatively less in counties with

²³The Social Capital Community Benchmark Survey was administered by the Saguaro Seminar at Harvard University. The Survey included a measure of informal social influence, which was constructed by principal components analysis of measures including having friends visiting home, visiting relatives, socializing with co-workers outside of work, hanging out with friends in public places, and playing cards and board games. Hence, it indirectly measures the opportunities for peer communication. Agarwal et al. (2009) used data from the same survey to analyze the impact of social interaction on the digital divide. Nevertheless, because the informal social influence variable was constructed by statistical procedures and survey responses, there could exist substantial measurement errors. Accordingly, we generated a binary index of social interaction by classifying informal social influence as high or low relative to the median. This is a more conservative way of using the social interaction data.

relatively high telemarketing sales.

Next, the county population might also affect DNC registrations through fixed costs in telemarketing. To the extent that telemarketing involves a fixed cost in each market, the business would be economic only if the consumer demand in that market is large enough. So, if sufficiently many consumers opt out, the demand for telemarketing might fall below the economic scale, and then the telemarketer would withdraw from the market. In this case, after September 2, when telemarketers began accessing the DNC registry, they would reduce solicitations to smaller markets and so fewer consumers would register with the DNC. This implies that the coefficient of number of households interacted with the indicator of September 2 and after would be *positive*.

The fixed costs of telemarketing would be most pertinent for niche demand such as from small consumer segments. In the contexts of radio and newspapers, George and Waldfogel (2003) and Waldfogel (2003) have found evidence of “preference externalities” with respect to race – it is economic for vendors to serve a racial segment only if it is sufficiently large. Accordingly, Table 3, column (5), reports an estimate including racial heterogeneity multiplied by the indicator of September 2 and after as an additional covariate. We computed racial heterogeneity as the probability that any two individuals drawn at random from a county would not belong to the same race. The coefficient of the number of households interacted with the time indicator of September 2 and after was positive and significant, albeit somewhat smaller than the main estimate. The coefficient of racial heterogeneity multiplied by the indicator of September 2 and after was positive and significant, suggesting that, on and after September 2, DNC registrations increased relatively *more* in racially heterogeneous counties, which is evidence against preference externalities.

Our last robustness test controlled for consumers’ English proficiency. Telemarketing is conducted by voice (though not necessarily human). So, in counties with more linguistically isolated consumers, who cannot communicate in English, the DNC registration rate could be high initially, but it might have changed when telemarketers accessed the registry from September 2 and revised their practices. To the extent that the county population was related to the language ability of consumers, the interaction of the number of households interacted with the indicator of September 2 and after could have represented vendor response to differences in language ability. To check this, Table 3, column (6), reports an estimate including the percent linguistically isolated multiplied by the indicator of September 2 and after as an additional covariate. The coefficient of the number of households interacted with the indicator of September 2 and after was positive and significant, and larger than the main estimate. The interaction of the percent linguistically isolated multiplied by the indicator of September 2 and after was not significant.

In general, the results of the robustness tests were consistent with the main estimate. Apparently, our finding that consumer DNC registration exhibited a positive externality from September 2 and after was robust to alternative specifications and interpretations, and possible confounding factors.

5.3 Falsifications

Besides checking for robustness to confounds and alternative explanations, another way to check our empirical findings is through falsification tests. These include placebo tests with alternative explanatory and dependent variables, and replicating the analysis in another setting where externality is minimal/does not exist. Table 4 reports the falsification analyses. For convenient reference, Table 4, column (1), reproduces the main estimate from Table 3, column (2).

First, recall that 27 states had established state-level DNC lists before the FTC set up the federal registry. As of September 2, 2003, 13 of these states had not merged their DNC lists with the federal registry (Varian et al. 2004). Consumers in these states would have experienced externalities after September 2 only to the extent that those who had not registered with the state DNC list did register with the federal DNC registry. So, the externality in these states should be weaker. Indeed, as Table 4, column (2), reports, among these states, the interaction of the number of households with the indicator of September 2 and after was positive but insignificant.

Next, we explored if the interaction of the number of households with the indicator of September 2 and after affected newspaper coverage of DNC. If it did, then there could be some unobserved factors around September 2 that also interacted with population to affect the broader market. These would confound our inference that externalities due to consumer and seller strategic behavior had affected DNC registrations. Table 4, column (3), reports an estimate with the number of newspaper reports of DNC per household, weighted by circulation in the county, as the dependent variable. The interaction of the number of households with the indicator of September 2 and after was not significant. Hence, our finding of externalities was specific to DNC registration.

Finally, we set up tests around other cut-off dates besides September 2. If we found that the interaction of the number of households with indicator of other cut-off dates was significant, then it would falsify our conclusion that the increased DNC registration was due to changes in seller solicitations after September 2. Table 4, columns (4) to (6), report regressions including number of households interacted with indicators of August 2 and

Table 4: Falsifications

VARIABLES	(1) September 2	(2) Unmerged State Lists	(3) News Reports	(4) Aug 2	(5) Oct 2	(6) Nov 2
Cumulative Registrations [t-1]	-0.065*** (0.008)	-0.028*** (0.006)	0.001 (0.001)	-0.011** (0.005)	-0.244*** (0.017)	-0.218*** (0.012)
Cum reg \times <i>Date</i>	-0.025*** (0.002)	-0.030*** (0.002)	-0.000 (0.000)	-0.001** (0.000)	-0.001*** (0.000)	0.001*** (0.000)
No. of households \times <i>Date</i>	0.267*** (0.067)	0.035 (0.056)	-0.029* (0.018)	-0.032 (0.020)	-0.004 (0.011)	-0.015 (0.009)
News reports	0.014 (0.017)	-0.027 (0.021)		-0.003 (0.014)	0.016 (0.011)	0.004** (0.002)
Constant	19.803*** (2.344)	3.153** (1.226)	-0.257 (0.210)	4.236*** (1.121)	73.416*** (5.143)	68.402*** (3.620)
Observations	16,506	11,340	26,586	16,506	16,506	16,506
R-squared	0.816	0.760	0.011	0.359	0.642	0.254
Number of fips	1,179	810	1,899	1,179	1,179	1,179

Notes. The dependent variable is log registration rate except for column (3). For ease of presenting the estimates, we rescaled log registration rate, log cumulative registration rate, and log news reports per household by multiplying them by 1000. All specifications included the interactions between *Date* and household size, household income, unemployment, commute time, and retail density as control variables (not reported), and county and day fixed effects, where *Date* refers to the indicator of *Date* and after as specified in the corresponding column headings. Robust standard errors clustered by county in parentheses. Column (1): Main estimate as reported in Table 3, column (2). Column (2): Excluding counties without state lists or with state lists that had been merged to the federal registry. Column (3): Dependent variable is number of newspaper reports of DNC per household weighted by circulation in the county. Column (4): Using data before/after August 2. Column (5): Using data before/after October 2. Column (6): Using data before/after November 2.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

after, October 2 and after, and November 2 and after (as with the main estimate in Table 3, the samples were limited to 7 days before/after the respective cut-off dates). All of these interactions were statistically insignificant, and, indeed, their coefficients were negative. Hence, we observed increased DNC registrations in more populated counties only after September 2, but not the other dates.

6 Implications and Conclusions

Our main empirical findings were that opt out facilities result in externalities from vendors to consumers, and implicitly, that consumers opt out and vendors adjust solicitations strategically according to population size. These findings provide important insights and

guidance to both managerial practice and public policy.

With the introduction of an opt out facility, viz., a DNC registry, consumers opt out according to their own preferences and the opt out decisions of others. The decisions of marginal consumers to opt out directly reduce the list of consumers available for solicitation. Since it is the marginal consumers, with relatively low benefit, who opt out, their opting out may actually help vendors by refining the list of consumers available for solicitation in favor of higher-benefit consumers. So, initially, the opt out facility may raise vendors' yield from solicitations. However, vendors may respond by re-directing their solicitations to consumers still available for solicitation. These consumers would experience an increase in solicitations, and some of them respond by opting out as well.

Hence, managers need to distinguish the initial impact of the opt out facility, which raises their yield from solicitations, from the long-term impact, which lowers their yield. Managers should plan according to the long-term equilibrium, in which consumers and vendors have fully adjusted to the opt-out facility.

Further, our empirical results show that the impact of an opt-out facility increases with the market size. The strategic behavior of consumers and vendors is relatively stronger in markets with bigger population. Apparently, the resulting externalities give rise to a *diseconomy of scale*. The larger is the market, the stronger are the externalities, and so, the larger would be the reduction in the long-term yield from solicitations.

Managers must balance this diseconomy of scale arising from externalities against any economy of scale in direct marketing. Direct marketing is attractive because of its apparent scalability – simply acquire a customer list and then use technology (and, possibly, outsourced labor) to solicit consumers. Our findings show that managers need to consider the countervailing effect due to externalities in consumers' opt out decisions.

Having chalked up millions of registrations over a short period of time, the federal DNC registry was widely hailed as a hugely popular government service (Taylor 2004). However, our results suggest caution from the perspective of public policy. After the initial burst of registrations, part of the subsequent registrations were driven by externalities rather than the consumers' own preferences. As other consumers opted out, vendors shifted solicitations, and so, led more consumers to opt out. Some of these consumers may indeed be interested in the marketed items.

Further, the impact of opt out varied according to population size. Here again, some consumers in a large market may be interested in the marketed items, but they opted out because they suffered additional harm from the mere size of their market (which caused

vendors to send more solicitations).

So, opt-out facilities give rise to complicated strategic interactions between consumers and vendors. To this extent, policy-makers should give more consideration to directly addressing the harm caused by solicitations through an appropriate tax. By directly addressing the harm, this policy would achieve more predictable and precise results, closer to economic efficiency.

Our empirical analysis was based on registrations with the federal DNC registry. It is important to ask whether the findings extend to other marketing contexts where consumers are concerned about privacy and where they could be provided with opt out facilities. Currently, a context of particular concern is online tracking over the Internet.

We do believe that the broad conclusions would apply to the online context. Provided with an opt out facility, if some consumers opt out, then they shrink the pool available to marketers. To the extent that the market size is large, vendors would spend more efforts in marketing, and some of these efforts would be re-directed to the pool of remaining consumers. These are the essential ingredients of the externality.

Accordingly, we expect a facility for opt out of tracking to generate externalities which increases with the consumer market size. The managerial and policy implications that we set out above should also apply.

Our results are subject to two major limitations. First, our theoretical model is static. It does not allow sellers to collect consumer information in one period and use it for marketing in a subsequent period. The model also does not consider repeat purchases. Such dynamic interactions, in the context of negative externalities imposed by seller solicitations, are an important direction for future work.

Second, while we employed two distinct empirical strategies, using instrumental variables in one and difference-in-differences in a natural experiment in the other, we could not absolutely rule out the relation between DNC registration and consumer population being due to some other unobserved factor. A particular gap was the absence of data on vendor solicitations. We could only infer vendor strategies from consumer DNC registrations. In future studies, it would be helpful to directly test the impact of consumer opt out on vendor marketing, so obtain more conclusive results.

References

- Agarwal, R., A. Animesh and K. Prasad, "Social Interactions and the "Digital Divide": Explaining Variations in Internet Use," *Information Systems Research*, Vol. 20 No. 2, June 2009, 277-294.
- Alesina, Alberto, and Eliana La Ferrara, "Participation in Heterogeneous Communities," *Quarterly Journal of Economics*, Vol. 115, No. 3, 2000, 847-904.
- Angrist, Joshua D., and Alan B. Krueger, "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, Vol. 15, No. 4, 2001, 69-85.
- Armstrong, Mark, John Vickers, and Jidong Zhou, "Consumer Protection and the Incentive to Become Informed," *Journal of the European Economic Association*, Vol. 7, Nos. 2-3, 2009, 399-410.
- Bertrand, Marianne, Erzo F.P. Luttmer, and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, Vol. 115, No. 3, 2000, 1019-1055.
- Bonfrer, Andre, and Xavier Dreze, "Real-Time Evaluation of E-mail Campaign Performance," *Marketing Science*, Vol. 28, No. 2, 2009, 251-263.
- Direct Marketing Association. *Power of Direct Marketing*, 2006-2007 Edition, 2006.
- Echambadi, Raj, and James D. Hess, "Mean-Centering Does Not Alleviate Collinearity Problems in Moderated Multiple Regression Models," *Marketing Science*, Vol. 26, No. 3, May-June 2007, 438-445.
- George, L. and J. Waldfogel, "Who Affects Whom in Daily Newspaper Markets?" *Journal of Political Economy*, Vol. 111, No. 4, 2003, 765-784.
- Goh, Khim-Yong, Kai-lung Hui, and I.P.L. Png, "Newspaper Reports and Consumer Choice: Evidence from the Do Not Call Registry," *Management Science*, forthcoming, 2011.
- Hann, Il-Horn, Kai-Lung Hui, Sang-Yong T. Lee, and I.P.L. Png, "Consumer Privacy and Marketing Avoidance: A Static Model," *Management Science*, Vol. 54 No. 6, June 2008, 1094-1103.
- Johnson, Justin P., "Targeted Advertising and Advertising Avoidance?" Johnson Graduate School of Management, Cornell University, November 7, 2008.
- Johnson, Peter A., "The New Economics of Telemarketing: Before and After the Federal DNC List," White Paper, Direct Marketing Association, New York, November 20, 2003.
- Mahajan, V., E. Muller, and F.M. Bass, "New Product Diffusion Models in Marketing: A Review and Directions for Research," *Journal of Marketing*, Vol. 54, No. 1, 1990, 1-26.
- Manski, Charles, "Identification of endogenous social effects: the reflection problem," *Review of Economic Studies*, Vol. 60, No. 3, 1993, 531-542.
- Marmaros, David and Bruce Sacerdote, "How Do Friendships Form," *Quarterly Journal of Economics*, Vol. 121, No. 1, 2006, 79-119.

- Milgrom, Paul and Chris Shannon, "Monotone Comparative Statics," *Econometrica*, Vol. 62, No. 1, 1994, 157-180.
- Rotfeld, Herbert Jack, "Do-not-call as the US Government's improvement to telemarketing efficiency," *Journal of Consumer Marketing*, Vol. 21, No. 4, 2004, 242-244.
- Staiger, Douglas and James H. Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica*, Vol. 65, No. 3, 1997, 557-586.
- Stock, James H. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg*, eds. J.H. Stock and D.W.K. Andrews, Cambridge University Press, 2005, 80-108.
- Taylor, Humphrey, "Do Not Call Registry Is Working Well", *Harris Interactive*, February 13, 2004.
- Tucker, Catherine, "Identifying Formal and Informal Influence in Technology Adoption with Network Externalities," *Management Science*, Vol. 54, No. 12, 2008, 2024-2038.
- U.S. Federal Trade Commission (FTC), "The Status of the National Do Not Call Registry," Prepared Statement before the Committee On Commerce, Science And Transportation, U.S. Senate, Washington, D.C., September 30, 2003.
- van den Berg, Gerard, "An Economic Analysis of Exclusion Restrictions for Instrumental Variable Estimation," Unpublished Manuscript, Princeton University, 2006.
- Van den Bulte, C., and S. Stremersch, "Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test," *Marketing Science*, Vol. 23, No. 4, 2004, 530-544.
- Van Zandt, Timothy, "An Introduction to Monotone Comparative Statics", Working Paper, INSEAD, November 2002.
- Varian, Hal, Fredrik Wallenberg, and Glenn Woroch, "Who Signed Up for the Do-Not-Call List?" School of Information, University of California, Berkeley, June 15, 2004.
- Varian, Hal, Fredrik Wallenberg, and Glenn Woroch, "The Demographics of the Do-Not-Call List," *IEEE Security & Privacy*, Vol. 3, 2005, 34-39.
- Waldfoegel, J., "Preference externalities: an empirical study of who benefits whom in differentiated-product markets," *RAND Journal of Economics*, Vol. 34, No. 3, 2003, 557-568.
- Wooldridge, J. M., "Score diagnostics for linear models estimated by two stage least squares," in *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao*, eds. G. S. Maddala, P. C. B. Phillips, and T. N. Srinivasan, Oxford: Blackwell, 1995, 66-87.
- Wooldridge, J.M. *Introductory Econometrics: A Modern Approach*. 3d edition. Thomson Southwestern, Mason, OH, 2006.

Appendix

Proof of Proposition 1.

By (1) and (2),

$$r_i^* = \arg \max_{r_i \in \{0,1\}} \left\{ [1 - r_i] \sum_{m=1}^M \frac{S_m}{n} [b_{im} - w] - r_i e \right\}.$$

Referring to Van Zandt (2002, Remark 3), we can ignore the term, $-r_i e$, as it depends on only r_i and not S_m .

Hence, the essential part of the maximand is

$$[1 - r_i] \sum_{m=1}^M \frac{S_m}{n} [b_{im} - w], \quad (\text{A1})$$

which simplifies to the product of two terms. One term, $1 - r_i$, is strictly decreasing in r_i , whereas the other is

$$\sum_{m=1}^M \frac{S_m}{n} [b_{im} - w]. \quad (\text{A2})$$

Now, $b_{\lambda m} < w$, and so, for the marginal consumer, $i = \lambda$, (A2) is strictly decreasing in S_m , strictly increasing in n , and strictly decreasing in w .

(i) The function, (A1), comprises the product of a term that is strictly decreasing in r_λ and another term that is strictly decreasing in S_m . So, by Van Zandt (2002, Proposition 1), the function, (A1), has strictly increasing differences in r_λ and S_m , $m = 1, \dots, M$, and so it is supermodular. By Van Zandt (2002, Theorem 1), the reaction function, r_λ^* , is increasing in S_m , and so, λ must increase in S_m .

(ii) The function, (A1), comprises the product of a term that is strictly increasing in $-r_\lambda$ and another term that is strictly increasing in n . Hence, the function, (A1), has strictly increasing differences in $-r_\lambda$ and n , and so it is supermodular. Accordingly, $-r_\lambda^*$ is increasing in n , and so, r_λ^* and hence λ are decreasing in n .

(iii) The function, (A1), comprises the product of a term that is strictly decreasing in r_λ and another term that is strictly decreasing in w . So, the function, (A1), has strictly increasing differences in r_λ and w , and so it is supermodular. Thus, r_λ^* is increasing in w , and so, λ is increasing in w .

Proof of Proposition 2.

(i) Referring to (4), we claim that, if $v_{\lambda m} \geq p$, then the function,

$$R(S_m, n) = \frac{S_m}{n} \sum_{i=\lambda+1}^N pq_{im}, \quad (\text{A3})$$

has increasing differences in S_m and n . By (3), $n = N - \lambda$. Hence, substituting from (A3),

$$\begin{aligned} & R(S_m + 1, n + 1) - R(S_m, n + 1) - R(S_m + 1, n) + R(S_m, n) \\ &= \frac{1}{n + 1} \sum_{i=\lambda}^N pq_{im} - \frac{1}{n} \sum_{i=\lambda+1}^N pq_{im} \geq 0 \end{aligned} \quad (\text{A4})$$

if and only if

$$\frac{\sum_{i=\lambda}^N q_{im}}{\sum_{i=\lambda+1}^N q_{im}} \geq \frac{n + 1}{n},$$

or

$$\frac{q_{\lambda m}}{\sum_{i=\lambda+1}^N q_{im}} \geq \frac{1}{n},$$

which would hold if $v_{\lambda m} \geq p$ because $\sum_{i=\lambda+1}^N q_{im} \leq n$. Accordingly,

$$R(S_m + 1, n + 1) - R(S_m, n + 1) \geq R(S_m + 1, n) - R(S_m, n), \quad (\text{A5})$$

and so $R(S_m, n)$ has increasing differences in S_m and n .

Consider the reaction function, S_m^* . By Van Zandt (2002, Theorem 5), if n increases, we must have either $S_m^*(n + 1) \geq S_m^*(n)$, or both $S_m^*(n + 1)$ and $S_m^*(n)$ maximize seller m 's profit. Hence, the reaction function, S_m^* , is weakly increasing in n .

(ii) Obvious, since the seller's profit does not vary with w .