Online Appendix to:

Price Cap vs. Ad Valorem Subsidies: Selection, Pricing, and Cross-Subsidization in the FCC's Rural Health Care Program

Dubey ⓡ Rabbani ⓡ Pinto

Author order randomized using the AEA Author Randomization Tool.

Not for typesetting. To be posted as supplemental material.

A Detailed description of the institutional environment

A.1 Internet subsidy programs in the US

The Federal Radio Commission (FRC) was the original government agency that regulated radio communication in the US from 1927 to 1934.[1] The Communications Act of 1934 abolished FRC and led to the creation of the Federal Communications Commission (FCC). Among other roles, the FCC provided subsidies to ensure that "*all people in the United States shall have access to rapid, efficient, nationwide communications service with adequate facilities at reasonable charges*".[2] The Telecommunications Act of 1996 substantially expanded the subsidy coverage, creating four sister programs, each targeting one recipient group as shown in Figure 1. The High Cost Program partially reimburses telecommunication companies that serve high-cost areas. E-Rate subsidizes internet access for schools and libraries. The Lifeline Program assists low-income consumers with their monthly telephone and telecommunication bills. The Rural Health Care Program (RHC) (the program that we study) subsidizes internet access and related equipment for eligible health care providers (HCPs).

The programs are funded and administered under the Universal Service Fund (USF). The FCC has appointed Universal Service Administrative Company (USAC), a private non-profit entity, to administer USF, while the FCC retains the right to oversight and assessment. USF is not dependent on the federal budget. Instead, funds are raised by taxing telecommunication companies through the "contribution factor." The contribution factor is the percentage of interstate end-user revenues that telecommunication companies must pay to USAC to fund USF. A complex system of accounting determines the contribution factor that each telecommunication company must pay.[3] USAC sets the contribution factor on a quarterly basis to ensure that program costs remain covered. Since 1997, USAC has set the contribution factor over 100 times.[4] While the FCC has the authority to appeal or adjust the contribution factor, it has never done so (Dunstan, 2023). Therefore, USAC is the de facto setter and enforcer of the contribution factor.

In response to the ever-increasing program expenditure, USAC has been steadily raising the contribution factor. Telecommunication companies are allowed to shift the contribution factor to consumers, itemized separately under "Universal Service Charge", "USF fee", or similar names.[5] Therefore, the contribution factor is a tax on ordinary internet users. The Government Accountability Office has "*raised concerns about . . . the cost burden [the fund] imposes on consumers*" (GAO, 2012).

Figure A1 shows the trend of the contribution factor since the program's inception. It began at 3.14% in 1998 (FCC, 1998). Within two years, it rose to 5.7% (FCC, 2000). When the

[1] Wikipedia: Federal Radio Commission
[2] FCC: Universal Service Fund
[3] Some explanations may be found on the FCC website.
[4] Contribution Factor & Quarterly Filings - USF Management Support.
[5] 47 CFR Ch. I § 54.712, USAC: Universal Service

contribution factor reached 16.1% in 2014, an FCC commissioner called it a "*clearly disturbing and unsustainable*" path.[6] This concern was echoed in the Senate in 2023 when the contribution factor reached 29%, where Senator John Thune emphasized the need to "*address the inefficiencies within*" each of the USF programs, pointing to documented waste, fraud, and abuse across them (U.S. Senate Subcommittee on Communications, Media, and Broadband, 2025, p.3-4). As of the first quarter of 2026, the contribution factor is 37.6%.[7]

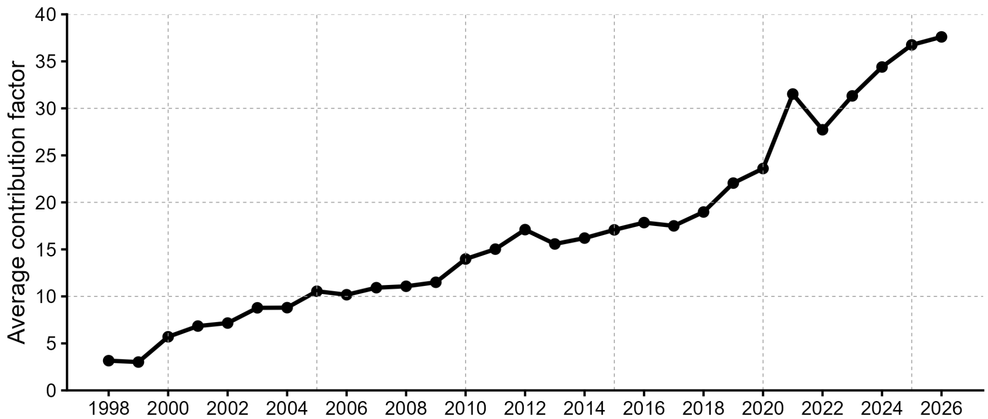

Figure A1: The evolution of the contribution factor.

Notes: Quarterly contribution-factor filings (FCC USAC) averaged to a single annual observation. Vertical axis is in percentage points.

Below is a back-of-the-envelope calculation of the effect of the program on internet prices in the US. The total Internet Service Provider (ISP) revenues in the US in 2025 are estimated to be $127.5 billion. The USF budget has been $8.6 billion in 2024 (USAC, 2024). Knowing that internet prices must be raised to fund USF, internet prices are likely inflated by $8.6/(127.5 - 8.6) \times 100 = 7.2\%$ via the contribution factor. The price elasticity of demand for residential internet is -0.18 to -0.47 (Czajkowski et al., 2024). Using this, the contribution factor likely priced out 1.3% to 3.4% of users, who could become reliant on subsidies to stay connected. This creates a positive feedback loop. Rising program costs increase the contribution factor, which raises internet prices, exacerbates affordability, increases internet users' subsidy reliance, and drives up subsidy expenditure (Hazlett, Schwall, and Wallsten, 2019).

It is understood in public choice theory that bureaucratic entities may seek to maximize the budget under their management as a means of political influence and financial power (Niskanen, 1968). Besides, USAC's own expenses are paid using the contribution factor (GAO, 2020), which amounted to $324 million in 2024 (USAC, 2024). Therefore, the relationship between the FCC and USAC may create a classical principal-agent problem, where cost-containment is not aligned with, or potentially contradicts, USAC's priorities. In a review of the High Cost program's 2008 spending, the FCC Office of Inspector General estimated that 23.3% of the subsidies were paid out "erroneously" (Wallsten, 2011, p.9), shedding light on the possible size of waste, fraud, and

[6] USF Contribution Factor Over Time.

[7] FCC:DA-25-223A1.

abuse. Despite the FCC's authority and central role, it has admitted to falling short in auditing cost information of entities under its oversight (Zolnierek, 2008).

A.2 The Rural Health Care program

RHC's original goal in 1997 was to ensure that rural HCPs have access to "telecommunications services necessary for the provision of health care services" (Gilroy, 2011, p.4), and that rural HCPs pay no more than their urban counterparts on internet bills.[8] This led to the creation of the first subsidy mechanism under the RHC, called the Telecommunications Program ($\mathcal{P}_1$) in 1997. $\mathcal{P}_1$ uses the following flat-rate subsidy reimbursement scheme: If an eligible rural HCP who pays internet price p could show that a similar plan costs p_u in a nearby urban location, the HCP would pay p_u out of its own pocket (HCP net cost), and the subsidy would cover $p - p_u$.

This design is economically problematic. If an HCP chooses an unjustifiably expensive internet plan, $\mathcal{P}_1$ fully compensates the HCP for the excessive payment. In contrast, if an HCP switches to a more cost-effective internet plan and minimizes waste, $\mathcal{P}_1$ fully captures the cost savings. This full cost insulation creates a perfectly inelastic demand for internet services, eliminating any cost-saving incentives (Rabbani, 2024a).

The Healthcare Connect Fund (HCF) was introduced in 2012, piloted in 2013, and fully implemented in 2014 to run alongside $\mathcal{P}_1$. Instead of benchmarking on urban rates, HCF reimburses HCPs for 65% of their rural internet bills while requiring the HCPs to pay the remaining 35%. This proportional cost-sharing incentivizes waste reduction because it rewards HCPs $0.35 for every $1.00 they save on their internet bills, and penalizes them $0.35 for every $1.00 they waste. Although none of the official FCC documents discloses the reasons for introducing HCF to run alongside $\mathcal{P}_1$, the FCC stated that they "expect" a gradual migration of $\mathcal{P}_1$ users to HCF[9], hinting that the introduction of HCF was an attempt to phase out $\mathcal{P}_1$'s suboptimal incentive design.

HCF allows HCPs to either individually request subsidies ($\mathcal{P}_2$) or to create a consortium of HCPs that requests subsidies on behalf of its members ($\mathcal{P}_2^c$). $\mathcal{P}_2^c$ extends subsidy coverage to eligible urban HCPs that belong to a majority-rural consortium. A consortium is deemed majority rural if more than half of its members are rural. It is based on a simple count of HCPs and not size or market capitalization. For example, if a large urban general hospital and two small rural clinics form a consortium, it is considered majority rural because two of the three members are rural, and all three members would be subsidized.

$\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_2^c$ use a common trio of eligibility criteria: Subsidy recipients must (1) be a non-profit or public entity, (2) be in a rural area or belong to a majority-rural consortium, and (3) belong to one of the following categories: non-profit hospitals, skilled nursing facilities, community mental health centers, rural health clinics, dedicated emergency departments of rural

[8] The FCC's Universal Service Rural Health Care Programs

[9] Healthcare Connect Fund - Frequently Asked Questions

for-profit hospitals, community health centers or health centers providing health care to migrants, local health departments or agencies, post-secondary educational institutions offering health care instruction such as teaching hospitals or medical schools, part-time eligible entities located in a facility that is ineligible, or consortia of HCPs consisting of one or more entities listed above.[10] Once eligibility is established, HCPs can choose either program as their subsidy channel, except for eligible urban HCPs who must remain in $\mathcal{P}_2^c$. HCPs can switch programs at the time of subsidy renewal.

Besides eligible rural and eligible urban, HCPs could be ineligible rural or ineligible urban. Ineligible HCPs do not count in the calculation that determines the majority-rural status of a consortium. When a consortium submits a subsidy request, it enters a competitive bidding process, in which ISPs compete by placing bids for providing all the internet services listed by the consortium for all its members. As explained in Section 1, we hypothesize that this could give rise to an unintended cross-subsidization scheme, where eligible HCPs covertly extend subsidy coverage to ineligible members in the same consortium. This could be done by artificially lowering the rates for ineligible members while raising the rates for eligible members in a manner that leaves the ISP revenue intact. For every $1 that is cross-subsidized, the consortium would collect an additional $0.65 in subsidies.

Table A1—HCP entity type composition by program and year, 2013–2014.

	2013	2014		
HCP type	$\mathcal{P}_1$	$\mathcal{P}_1$	$\mathcal{P}_2$	$\mathcal{P}_2^c$
Rural health clinic	262	136	96	21
Non-profit hospital	398	144	150	7
Mental health center	84	49	18	0
Community health center	96	67	24	1
Local health department	115	107	2	0
Skilled nursing facility	0	0	0	0
Medical school	3	1	2	0
Emergency room	3	0	2	0
Not available	2	2	0	0

Notes: Cells report counts of HCPs (sample of 970 baseline HCPs). HCPs with mixed program participation within a year are excluded.

Table A1 shows the composition of HCPs that received subsidies under each program during 2013–2014. Non-profit hospitals make up the largest group, followed by rural health clinics, local health departments, community health centers, and mental health centers.

A.3 Institutional underpinnings of cross-subsidization

This section describes the institutional settings that may enable cross-subsidization. Any group of HCPs could consolidate all its subsidy requests into a single consortium application. The application must enter what resembles an auction process that allows ISPs to place bids to provide services

[10] https://www.usac.org/rural-health-care/telecommunications-program/step-1-determine-eligibility-of-your-site/

to the consortium. The process begins with HCPs filing Form 465 to USAC to officially declare and detail their internet connectivity needs (bandwidth, locations, technology, etc.) in a Request For Proposal (RFP). USAC puts RFPs on its website for 28 days. This is called the "competitive bidding" process, which is the ISPs' opportunity to place bids. After this window ends, the consortium can review the bids and select the winner. The consortium may sign a contract with the winner as early as on day 29, without USAC's approval, so long as the HCP can show that it has chosen the "most cost-effective" bid and document its evaluation criteria (USAC, 2025).

Table A2—RFP criteria examples.

	Criteria	Weight (%)
Consortium 1	Cost	16
	Overall network solution	14
	24 hour support team	13
	Prior experience including past performance	12
	Prices for ineligible services, products and fees	14
	Provisioning and installation time	11
	Leveraging existing technology	10
	Reliability of equipment (new v. refurbished)	10
Consortium 2	Cost	50
	Prior experience including past performance	50
Consortium 3	Prior experience, including past performance	20
	Reliability of Service	30
	Other (solicitation compliance)	10
	Cost	30
	Project management plan	10
Consortium 4	Cost	20
	Reliability of service	20
	Other (consideration of early termination fees)	15
	Leverage existing resources	15
	Prior experience, including past performance	15
	Technical support	15

Notes: Criteria and weights from four subsidized consortia's competitive-bid RFPs. Consortium identities are masked but all four are subsidy recipients.

Notably, the auction winner is not the ISP that places the lowest bid. Instead, the consortium is at liberty to assign any weights to any objective or subjective criteria that it deems to be important (FCC, 2014b, p.3). Table A2 shows four actual instances of criteria and weights that were used in RFPs. This allows consortia to engineer the criteria and weights to make a particular ISP win, even if the ISP's bid is objectively uncompetitive.

Even the legal language that is meant to prevent abuse seems vague and unenforceable. The RFP submission system requires the consortium representative to certify under penalty of perjury that they have selected the "*method that costs the least after consideration of the features, quality of transmission, reliability, and other factors that the health care provider deems relevant to choosing a method of providing the required health care services.*" (FCC, 2014c, p.2).

This process does not have the economic building blocks of an auction. An auction has quantitative metrics in place that are public information before the auction begins, so that a given set of bids results in a deterministically or probabilistically predictable outcome, a deviation from which could be appealed. In the case of RHC, the process is based on subjective and unquantifiable measures that the consortium sets and enforces, and its subjectivity makes it unobjectionable.

Instead, this process resembles a procurement mechanism with pure buyer discretion at the cost of a third party. This unusual combination (the consortium sets the auction criteria, selects the winner, and makes taxpayers pay the price) gives the consortium the utmost negotiating leverage. Therefore, the ISP's best response may not entail competitive bidding. Instead, an ISP's winning strategy may be to surpass other bidders in bending the terms of the trade in favor of the consortium. Of course, there is a limit to how much the terms of the trade could be distorted before it becomes unattractive to the bidding ISP.

We hypothesize that this process allows for devising a scheme in which a surplus is transferred from taxpayers to the consortium without cannibalizing the ISP's revenue. Each consortium may have members (HCPs) that are eligible to receive subsidies, and members that are ineligible. Eligibility is determined before competitive bidding begins (FCC, 2014a, p.5). So, when bids are being placed and negotiations are ongoing, both parties (consortium and ISP) know which lines of internet connection belong to eligible members and which lines belong to ineligible members. Both parties know that every \$1 of cross-subsidization transfers \$0.65 from taxpayers to the consortium in a way that is revenue-neutral to the ISP.

Given that each RFP has many details that may need clarification, consortia and ISPs are allowed to engage in one-on-one discussions to clarify needs and expectations before submitting bids. Moreover, USAC allows ISP who bid in some auctions to act as a consortium representative or consultant in other auctions, provided the ISP maintains sufficient organizational separation between staff involved in bidding and those engaged in consortium consulting activities (FCC, 2014a, p.10). This may facilitate collaboration between consortia and ISPs, weaken the competitive nature of the bidding process, and induce cartel-like behaviors.

If a consortium and ISP deem the bidding process as one stage of a repeated game, a cooperative equilibrium may be maintained in the long run. In this setting, collusion is lucrative and likely. In 2023, an ISP named GCI Communication Corp agreed to pay \$42.1 million to settle a fraud allegation case where it was accused of violating the competitive bidding regulations and colluding with an HCP to inflate prices and knowingly receive excessive subsidy payments during 2015-2018 (FCC order DA 23-380). The data shows that, during 2012-2023, GCI was the service provider to 4275 lines of subsidy, and it has received \$1,237,016,437 in subsidy payments. So, the \$42.1 million settlement claws back only 3.4% of the total subsidies GCI collected since 2012. This creates a clear economic incentive to engage in fraudulent behavior. The potential gains from inflating claims far exceed the financial risk posed by penalty and enforcement. In effect, the penalty functions as a minor operational expense rather than a meaningful deterrent, rendering the expected payoff of collusion overwhelmingly positive.

B Model proofs and derivations

B.1 Proof of Proposition P.1

When the cap binds ($p \geq \bar{p}$), the consumer price is pinned at $\bar{p}$ and demand is $D(\bar{p})$. The firm's objective (3) is

$$\pi(p) = (p - c)\, D(\bar{p}) - \alpha\, \Phi(p - \bar{p}).$$

The first-order condition with respect to p is

$$D(\bar{p}) - \alpha\, \Phi'(p - \bar{p}) = 0.$$

Since $\Phi'' > 0$ and $\Phi'(0) = 0$, the marginal penalty Φ' is strictly increasing from zero. Hence the equation $\Phi'(p - \bar{p}) = D(\bar{p})/\alpha$ has a unique solution

$$p^{cap} - \bar{p} = (\Phi')^{-1}\!\left(\frac{D(\bar{p})}{\alpha}\right),$$

establishing (4). Because $(\Phi')^{-1}$ is increasing, p^{cap} is increasing in $D(\bar{p})$ and decreasing in α. The second-order condition $-\alpha\, \Phi''(p - \bar{p}) < 0$ holds by $\Phi'' > 0$, confirming a maximum.

Under the quadratic penalty $\Phi(\delta) = \frac{\gamma}{2}\delta^2$, $\Phi'(\delta) = \gamma\delta$, so $p^{cap} - \bar{p} = D(\bar{p})/(\alpha\gamma)$. Government outlays are

$$G^{cap} = (p^{cap} - \bar{p})\, D(\bar{p}) = \frac{D(\bar{p})}{\alpha\gamma} \cdot D(\bar{p}) = \frac{D(\bar{p})^2}{\alpha\gamma}. \qquad \square$$

B.2 Proof of Proposition P.2

Part (i). By the effective-marginal-cost representation (8), $p_c^{adv}(\tau) = p^{no}\big(c(1-\tau)\big)$. By the regularity condition in Section 3.3, the map $\tau \mapsto p_c^{adv}(\tau)$ is continuous and strictly decreasing. Moreover, condition (9) gives $p_c^{adv}(0) = p^{no}(c) > \bar{p}$ and $\lim_{\tau\uparrow 1} p_c^{adv}(\tau) = p^{no}(0) < \bar{p}$. Therefore there is a unique $\tau^* \in (0,1)$ satisfying $p_c^{adv}(\tau^*) = \bar{p}$ (10). For $\tau \geq \tau^*$, monotonicity gives $p_c^{adv}(\tau) \leq p_c^{adv}(\tau^*) = \bar{p}$, with strict inequality when $\tau > \tau^*$. Since the cap binds, $p_c^{cap} = \bar{p}$.

Part (ii). Since $D' < 0$ and $p_c^{adv}(\tau) \leq \bar{p} = p_c^{cap}$ by Part (i),

$$Q^{adv}(\tau) = D\big(p_c^{adv}(\tau)\big) \geq D(\bar{p}) = Q^{cap}.$$

Part (iii). Define $E(p) \equiv p\, D(p)$. Then

$$E'(p) = D(p) + p\, D'(p) = D(p)\big(1 - \varepsilon_D(p)\big), \tag{1}$$

where $\varepsilon_D(p) \equiv -pD'(p)/D(p)$ is the price elasticity of demand. If $\varepsilon_D(p) \leq 1$ on $[p_c^{adv}(\tau), \bar{p}]$, then $E'(p) \geq 0$ on that interval, so E is nondecreasing. Since $p_c^{adv}(\tau) \leq \bar{p}$,

$$E^{adv}(\tau) = E\big(p_c^{adv}(\tau)\big) \leq E(\bar{p}) = E^{cap}.$$

Part (iv). Under the quadratic penalty $\Phi(\delta) = \frac{\gamma}{2}\delta^2$, Proposition P.1 gives $G^{cap} = D(\bar{p})^2/(\alpha\gamma)$. For any fixed $\tau \in (0,1)$, the ad valorem equilibrium price $p^{adv}(\tau)$ is finite, since it solves the smooth first-order condition (6) with bounded primitives. Therefore, $G^{adv}(\tau) = \tau\, p^{adv}(\tau)\, Q^{adv}(\tau)$

is finite and independent of (α, γ). As $\alpha\gamma \to 0$, $G^{cap} \to \infty$ while $G^{adv}(\tau)$ remains bounded. Hence $G^{adv}(\tau) < G^{cap}$ for sufficiently small $\alpha\gamma$. □

B.3 Proof of Proposition P.3

From (14), the consortium's reduced-form objective is

$$\Psi(\kappa) = B\,\kappa - \alpha\,\Phi\big(B(\kappa - 1);\, R\big), \qquad \kappa \in [1,\, 1 + R].$$

The proposition treats the nondegenerate case $B > 0$ and $R > 0$. If $R = 0$, the feasible set is the singleton $\{1\}$, so $\kappa^* = 1$ and no cross-subsidization is feasible.

Part (i). Differentiating, $\Psi'(\kappa) = B - \alpha\, B\, \partial\Phi/\partial\delta\big(B(\kappa-1);\, R\big) = B\big[1 - \alpha\, \partial\Phi/\partial\delta\big(B(\kappa-1);\, R\big)\big]$. Evaluating at $\kappa = 1$:

$$\Psi'(1) = B\big[1 - \alpha\, \partial\Phi/\partial\delta(0;\, R)\big] = B > 0,$$

where the last equality uses condition (ii), $\partial\Phi/\partial\delta(0;\, R) = 0$. Since $R > 0$, values $\kappa = 1 + \varepsilon$ are feasible for sufficiently small $\varepsilon > 0$. Because $\Psi'(1) > 0$, such a small increase in κ raises the consortium's payoff, so $\kappa^* > 1$ and $\tilde{p}_E > p_E^{adv}$.

Part (ii). The second derivative is

$$\Psi''(\kappa) = -\alpha\, B^2\, \partial^2\Phi/\partial\delta^2\big(B(\kappa - 1);\, R\big) < 0 \qquad \text{for all } \kappa \geq 1,$$

by condition (iv), $\partial^2\Phi/\partial\delta^2 > 0$. Hence Ψ is strictly concave on $[1,\, 1 + R]$, and the maximizer κ^* is unique.

Part (iii). By part (i), $\Psi'(1) > 0$, so $\kappa^* > 1$. At the upper boundary, $\Psi'(1 + R) = B\big[1 - \alpha\, \partial\Phi/\partial\delta(BR;\, R)\big]$. If $\alpha\, \partial\Phi/\partial\delta(BR;\, R) > 1$, then $\Psi'(1 + R) < 0$. Since Ψ' is continuous, strictly decreasing (by $\Psi'' < 0$), positive at $\kappa = 1$, and negative at $\kappa = 1 + R$, the intermediate value theorem yields a unique interior $\kappa^* \in (1,\, 1 + R)$ satisfying $\Psi'(\kappa^*) = 0$, which gives (15). If $\alpha\, \partial\Phi/\partial\delta(BR;\, R) \leq 1$, then $\Psi'(\kappa) \geq 0$ for all $\kappa \in [1,\, 1 + R]$, so Ψ is nondecreasing on the entire domain and $\kappa^* = 1 + R$.

Part (iv). For $R > 0$, under $\Phi(\delta;\, R) = \frac{\gamma R}{2}\, \delta^2$, verify: (i) $\Phi(0;\, R) = 0$; (ii) $\partial\Phi/\partial\delta(0;\, R) = \gamma R \cdot 0 = 0$; (iii) $\partial\Phi/\partial\delta(\delta;\, R) = \gamma R\, \delta > 0$ for $\delta > 0$; (iv) $\partial^2\Phi/\partial\delta^2(\delta;\, R) = \gamma R > 0$; (v) $\partial^2\Phi/\partial\delta\, \partial R = \gamma > 0$. The first-order condition (15) becomes $\alpha\, \gamma R\, B(\kappa^* - 1) = 1$, yielding $\kappa^* - 1 = 1/(\alpha\gamma B\, R)$. This is feasible ($\kappa^* - 1 \leq R$) if and only if $1/(\alpha\gamma B\, R) \leq R$, i.e., $R \geq R^* \equiv 1/\sqrt{\alpha\gamma B}$. When $R < R^*$, the interior solution exceeds the upper bound, so the constraint binds and $\kappa^* = 1 + R$. Therefore

$$\kappa^*(R) = \min\left\{1 + R,\ 1 + \frac{1}{\alpha\gamma B\, R}\right\}.$$

For $R \leq R^*$: $\kappa^*(R) = 1 + R$, strictly increasing. For $R \geq R^*$: $\kappa^*(R) = 1 + 1/(\alpha\gamma B\, R)$, strictly decreasing. Both branches meet at $R = R^*$ with $\kappa^*(R^*) = 1 + R^*$. As $R \downarrow 0$, $\kappa^* \to 1$. As $R \to \infty$, $\kappa^* \to 1$. Hence κ^* is hump-shaped, attaining its unique maximum at R^*. □

B.4 Closed-form solutions under linear demand

Throughout this subsection, specialize to linear demand $D(p_c) = a - bp_c$ with $a, b > 0$ and $a > bc$ (to ensure positive equilibrium quantities), and use the quadratic penalty $\Phi(\delta) = \frac{\gamma}{2}\delta^2$ ($\gamma > 0$).

No-subsidy monopoly ($M0$)

The monopolist solves $\max_{p \geq 0} (p - c)(a - bp)$. The first-order condition $a - bp - (p - c)b = 0$ yields

$$p^{no} = \frac{1}{2}\left(\frac{a}{b} + c\right), \qquad Q^{no} = \frac{1}{2}(a - bc). \tag{2}$$

Profits, consumer expenditures, and government outlays are

$$\pi^{no} = \frac{(a - bc)^2}{4b}, \qquad E^{no} = p^{no} Q^{no} = \frac{(a + bc)(a - bc)}{4b}, \qquad G^{no} = 0.$$

Price-cap regime

Under a binding cap ($p \geq \bar{p}$), demand is $D(\bar{p}) = a - b\bar{p}$. Proposition **P.1** with $\Phi'(\delta) = \gamma\delta$ gives

$$p^{cap} = \bar{p} + \frac{a - b\bar{p}}{\alpha\gamma}, \qquad G^{cap} = \frac{(a - b\bar{p})^2}{\alpha\gamma}. \tag{3}$$

The cap binds whenever the unconstrained monopoly price exceeds $\bar{p}$, i.e., $p^{no} > \bar{p}$, which requires $\bar{p} < \frac{1}{2}(a/b + c)$.

Ad valorem regime

The firm maximizes $(p - c)(a - b(1 - \tau)p)$. The first-order condition yields

$$p^{adv}(\tau) = \frac{1}{2}\left(\frac{a}{b(1 - \tau)} + c\right). \tag{4}$$

The consumer price, quantity, and government outlays are

$$p_c^{adv}(\tau) = (1 - \tau)\, p^{adv}(\tau) = \frac{1}{2}\left(\frac{a}{b} + c(1 - \tau)\right), \tag{5}$$

$$Q^{adv}(\tau) = a - b\, p_c^{adv}(\tau) = \frac{1}{2}\big(a - bc(1 - \tau)\big), \tag{6}$$

$$G^{adv}(\tau) = \tau\, p^{adv}(\tau)\, Q^{adv}(\tau). \tag{7}$$

Sufficient condition for ad valorem dominance

Condition (10) requires $p_c^{adv}(\tau) \leq \bar{p}$. Substituting (5):

$$\frac{1}{2}\left(\frac{a}{b} + c(1 - \tau)\right) \leq \bar{p} \quad \iff \quad \tau \geq 1 - \frac{2\bar{p} - a/b}{c}, \tag{8}$$

provided the right-hand side lies in $(0, 1)$.

Fiscal comparison under linear demand

From (3), $G^{cap} = (a - b\bar{p})^2/(\alpha\gamma)$. As $\alpha\gamma \to 0$, $G^{cap} \to \infty$. For any fixed $\tau \in (0, 1)$, $G^{adv}(\tau)$ is bounded, confirming Proposition **P.2**(iv).

To illustrate the comparison at finite $\alpha\gamma$, note that $G^{adv}(\tau) < G^{cap}$ whenever

$$\tau\, p^{adv}(\tau)\, Q^{adv}(\tau) < \frac{(a - b\bar{p})^2}{\alpha\gamma}.$$

Using (4)–(6), the left-hand side is a function of (τ, a, b, c) that does not depend on $\alpha\gamma$, while the right-hand side is unbounded as $\alpha\gamma \to 0$. Hence for any parameter configuration satisfying $a > bc$ and condition (8), there exists a threshold $\overline{\alpha\gamma} > 0$ below which $G^{adv}(\tau) < G^{cap}$.

Consortium under linear demand

With linear demand $D_j(p) = a_j - b_j p$ for each member $j \in \{E, I\}$, the monopoly equilibrium prices are

$$p_E^{adv} = \frac{1}{2}\left(\frac{a_E}{b_E(1-\tau)} + c\right), \qquad p_I^{adv} = \frac{1}{2}\left(\frac{a_I}{b_I} + c\right).$$

The reduced-form objects defined in (12) become

$$B = \tau\, p_E^{adv}\big(a_E - b_E(1-\tau)p_E^{adv}\big), \qquad R = \frac{p_I^{adv}(a_I - b_I p_I^{adv})}{p_E^{adv}(a_E - b_E(1-\tau)p_E^{adv})}.$$

Proposition **P.3** then delivers the optimal distortion

$$\kappa^*(R) = \min\left\{1 + R,\ 1 + \frac{1}{\alpha\gamma B\, R}\right\},$$

with peak distortion at $R^* = 1/\sqrt{\alpha\gamma B}$.

For the symmetric-demand case $D_E = D_I = D$ with $a_E = a_I = a$, $b_E = b_I = b$:

$$p_E^{adv} = \frac{1}{2}\left(\frac{a}{b(1-\tau)} + c\right), \qquad p_I^{adv} = \frac{1}{2}\left(\frac{a}{b} + c\right) = p^{no}.$$

The revenue ratio simplifies to

$$R = \frac{p^{no}\, D(p^{no})}{p_E^{adv}\, D\big((1-\tau)p_E^{adv}\big)},$$

which is a function of τ alone (for given a, b, c). As $\tau \to 0$, $p_E^{adv} \to p^{no}$ and $R \to 1$, while the subsidy base $B \to 0$, so $R^* \to \infty$ and $\kappa^* \to 1 + R$. The constraint binds but the subsidy base is negligible. To see the opposite limit, write $x \equiv 1 - \tau$. Then

$$B = \tau\,\frac{a^2 - b^2c^2x^2}{4bx}, \qquad R = \frac{(a^2 - b^2c^2)x}{a^2 - b^2c^2x^2}.$$

As $\tau \to 1$ (equivalently, $x \downarrow 0$), $B \to \infty$ and $R \to 0$. Since

$$\frac{R}{R^*} = R\sqrt{\alpha\gamma B} \to 0,$$

we have $R < R^*$ near the limit. Thus the feasibility constraint, not enforcement, binds, and $\kappa^* = 1 + R \to 1$.

C Program-share identification

The program-share specification in Section 4.3 uses continuous variables because the data are aggregated to the HCP level. The primitive treatment, however, is discrete: an underlying line, facility, or funded service is assigned to $\mathcal{P}_1$, $\mathcal{P}_2$, or $\mathcal{P}_2^c$. The variables $S_{jt,2}$ and $S_{jt,2^c}$ are therefore not continuous policy doses. They are HCP-level shares of subsidized activity assigned to $\mathcal{P}_2$ and $\mathcal{P}_2^c$ in period t.

This appendix states the interpretation of the first-differenced version of the continuous program-share specification in the main text. Let

$$\Delta \bar{Y}_j \equiv \bar{Y}_{j1} - \bar{Y}_{j0}$$

denote the change in the quantity-weighted average outcome for HCP j between $t = 0$ and $t = 1$, and let

$$\mathbf{S}_{j1} \equiv (S_{j1,2}, S_{j1,2^c})'$$

denote the post-period program-share vector. Since $S_{j0,d} = 0$ for $d \in \{2, 2^c\}$, the first-differenced share regression is

$$\Delta \bar{Y}_j = \lambda + \tau_{1,2} S_{j1,2} + \tau_{1,2^c} S_{j1,2^c} + u_j. \tag{9}$$

The same interpretation applies after residualizing both sides with respect to the included controls.

Let $\mathcal{I}_j$ denote the set of underlying lines or facilities in HCP j, $D_{i,1} \in \{1, 2, 2^c\}$ the 2014 program assignment at facility i (with $D_{i,0} = 1$ for all i), $Y_{i,t}(d)$ the facility-level potential outcome under program d in period t, and $Q_{i,t}$ the realized period-t facility quantity. For each destination program $d \in \{2, 2^c\}$ and each HCP with $S_{j1,d} > 0$, define the realized-quantity-weighted mean facility-level effect of switching from $\mathcal{P}_1$ to d as

$$\bar{\delta}_{j,d} \equiv \frac{\sum_{i \in \mathcal{I}_j} \mathbf{1}[D_{i,1} = d] \left[Y_{i,1}(d) - Y_{i,1}(1)\right] Q_{i,1}}{\sum_{i \in \mathcal{I}_j} \mathbf{1}[D_{i,1} = d]\, Q_{i,1}}; \tag{10}$$

the product $S_{j1,d}\, \bar{\delta}_{j,d}$ is taken to be zero when $S_{j1,d} = 0$. Let

$$\Delta_j \equiv \frac{\sum_{i \in \mathcal{I}_j} Y_{i,1}(1)\, Q_{i,1}}{\sum_{i \in \mathcal{I}_j} Q_{i,1}} - \frac{\sum_{i \in \mathcal{I}_j} Y_{i,0}(1)\, Q_{i,0}}{\sum_{i \in \mathcal{I}_j} Q_{i,0}}$$

denote the change in $\bar{Y}_j$ that would obtain at $t = 1$ if every facility in j remained on $\mathcal{P}_1$, aggregated with the realized post-period quantity weights. The following identity is mechanical under no within-HCP interference at the facility level: in period 1, each facility i assigned to d contributes $[Y_{i,1}(d) - Y_{i,1}(1)]\, Q_{i,1}$ to $\bar{Y}_{j,1} - \Delta_j - \bar{Y}_{j,0}$ beyond its untreated potential outcome, and the sum of these contributions over $i \in \mathcal{F}_{j,d} \equiv \{i : D_{i,1} = d\}$ collapses to $S_{j1,d}\bar{\delta}_{j,d}$ by (10). Hence

$$\Delta \bar{Y}_j = \Delta_j + S_{j1,2}\, \bar{\delta}_{j,2} + S_{j1,2^c}\, \bar{\delta}_{j,2^c}. \tag{11}$$

Define the activity-weighted average treatment effect on the treated (ATT) aggregate for destination d as

$$\tau^S_{1,d} \equiv \frac{\mathbb{E}[S_{j1,d}\bar{\delta}_{j,d}]}{\mathbb{E}[S_{j1,d}]}, \qquad d \in \{2, 2^c\}, \tag{12}$$

whenever $\mathbb{E}[S_{j1,d}] > 0$. This is the mean treatment effect per unit of post-period subsidized activity assigned to destination program d.

Assumptions. The HCP-level parallel-trends condition requires that, absent switching from $\mathcal{P}_1$, the expected change in the HCP-level quantity-weighted outcome would not vary with the post-period program-share vector, after conditioning on the included controls. The aggregation restriction is that the quantity-weighted mean facility-level effect of switching from $\mathcal{P}_1$ to destination program d is invariant to the HCP-level program-share vector, after conditioning on the included controls:

$$\mathbb{E}[\bar{\delta}_{j,d} \mid \mathbf{S}_{j1}] = \tau^S_{1,d}, \qquad d \in \{2, 2^c\}. \tag{13}$$

This restriction allows treatment effects to vary across facilities and HCPs. It rules out systematic correlation between treatment-effect heterogeneity and the share variation used to estimate equation (9).

Proposition P.1. Let $\mathbf{X}_j \equiv (1,\, S_{j1,2},\, S_{j1,2^c})'$. Suppose HCP-level parallel trends and the aggregation restriction in (13) hold, and that $\mathbb{E}[\mathbf{X}_j\mathbf{X}_j']$ is positive definite (ruling out, for instance, the boundary case $S_{j1,2} + S_{j1,2^c} \equiv 1$ in which no HCP remains on $\mathcal{P}_1$ at $t = 1$). Then the population coefficients on $S_{j1,2}$ and $S_{j1,2^c}$ in (9) equal the activity-weighted ATT aggregates:

$$\tau_{1,2} = \tau^S_{1,2}, \qquad \tau_{1,2^c} = \tau^S_{1,2^c}.$$

Proof. Taking conditional expectations of the decomposition (11) given $\mathbf{S}_{j1}$ (and the controls, suppressed in the notation; the residualization argument below is unchanged),

$$\mathbb{E}[\Delta\bar{Y}_j \mid \mathbf{S}_{j1}] = \mathbb{E}[\Delta_j \mid \mathbf{S}_{j1}] + S_{j1,2}\,\mathbb{E}[\bar{\delta}_{j,2} \mid \mathbf{S}_{j1}] + S_{j1,2^c}\,\mathbb{E}[\bar{\delta}_{j,2^c} \mid \mathbf{S}_{j1}].$$

Under HCP-level parallel trends, $\mathbb{E}[\Delta_j \mid \mathbf{S}_{j1}]$ does not depend on $\mathbf{S}_{j1}$; after residualizing the controls, write its common value as λ_0. Applying the aggregation restriction in (13) substitutes $\mathbb{E}[\bar{\delta}_{j,d} \mid \mathbf{S}_{j1}] = \tau^S_{1,d}$ to yield

$$\mathbb{E}[\Delta\bar{Y}_j \mid \mathbf{S}_{j1}] = \lambda_0 + S_{j1,2}\tau^S_{1,2} + S_{j1,2^c}\tau^S_{1,2^c}.$$

Thus the conditional expectation of $\Delta\bar{Y}_j$ is exactly linear in the regressors of (9). Because $\mathbb{E}[\mathbf{X}_j\mathbf{X}_j']$ is positive definite, the population linear projection coefficients equal the slopes of this conditional expectation. Hence $\tau_{1,2} = \tau^S_{1,2}$ and $\tau_{1,2^c} = \tau^S_{1,2^c}$. □

When the aggregation restriction fails. If HCP-level parallel trends holds but $\mathbb{E}[\bar{\delta}_{j,d} \mid \mathbf{S}_{j1}]$ varies with the share vector, then equation (9) no longer identifies the activity-weighted ATT aggregates in (12). The population coefficients remain well-defined linear projection coefficients of $\Delta\bar{Y}_j$ on the share vector:

$$(\tau^{LP}_{1,2}, \tau^{LP}_{1,2^c}) = \arg\min_{b_2, b_{2^c}} \mathbb{E}\Big[\big(\Delta\bar{Y}_j - \lambda - b_2 S_{j1,2} - b_{2^c} S_{j1,2^c}\big)^2\Big],$$

with the intercept optimized jointly. These coefficients summarize how HCP-level outcome changes covary with program shares under the maintained parallel-trends condition. They are not, without the aggregation restriction, nonparametric binary ATTs. In this sense the regression is a transparent linear summary of discrete underlying transitions rather than an application of a continuous-dose treatment model.

Relation to continuous-treatment DiD. Callaway, Goodman-Bacon, and Sant'Anna (2024) study difference-in-difference (DiD) designs in which the primitive treatment is genuinely continuous, and their central diagnostic—that linear regression on a continuous treatment recovers a structurally interpretable aggregate only under restrictions on how treatment effects vary with the regressor—applies here in the analogous form given by (13). The present setting differs from theirs in that the primitive assignment is discrete; the continuous regressors $S_{j1,2}$ and $S_{j1,2^c}$ arise from HCP-level aggregation rather than from a genuine dose.

D Double/Debiased machine learning implementation

This section details the implementation of the machine learning method. We estimate the following partially linear regression (PLR) model:

$$\bar{Y}_{jt} = \boldsymbol{S}_{jt}' \boldsymbol{\theta}_0 + g_0(\boldsymbol{X}_{jt}) + \varepsilon_{jt}, \qquad \mathbb{E}[\varepsilon_{jt} \mid \boldsymbol{X}_{jt}, \boldsymbol{S}_{jt}] = 0, \tag{14}$$

where $\boldsymbol{\theta}_0$ is the parameter vector of interest, $g_0(\boldsymbol{X}_{jt})$ is an unknown confounding function, and $\boldsymbol{X}_{jt} \in \mathbb{R}^p$ is a vector of covariates. The nuisance functions are:

$$\ell_0(\boldsymbol{x}) = \mathbb{E}[\bar{Y}_{jt} \mid \boldsymbol{X}_{jt} = \boldsymbol{x}], \qquad \boldsymbol{m}_0(\boldsymbol{x}) = \mathbb{E}[\boldsymbol{S}_{jt} \mid \boldsymbol{X}_{jt} = \boldsymbol{x}] \in \mathbb{R}^{N_d}, \tag{15}$$

where $\boldsymbol{X}_{jt}$ denotes the conditioning variables. The Double/Debiased Machine Learning (DML) estimator leverages the Neyman-orthogonal score for the PLR model: $\psi(\boldsymbol{W}_{jt}; \boldsymbol{\theta}, \eta) = \big(\boldsymbol{S}_{jt} - \boldsymbol{m}(\boldsymbol{X}_{jt})\big)\Big(\bar{Y}_{jt} - \ell(\boldsymbol{X}_{jt}) - \big(\boldsymbol{S}_{jt} - \boldsymbol{m}(\boldsymbol{X}_{jt})\big)'\boldsymbol{\theta}\Big)$, where $\boldsymbol{W}_{jt} = (\bar{Y}_{jt}, \boldsymbol{S}_{jt}, \boldsymbol{X}_{jt})$, $\eta = (\ell, \boldsymbol{m})$, and $\mathbb{E}[\psi(\boldsymbol{W}_{jt}; \boldsymbol{\theta}_0, \eta_0)] = \boldsymbol{0}$. The score satisfies Neyman orthogonality: $\frac{\partial}{\partial t}\mathbb{E}\left[\psi(\boldsymbol{W}_{jt}; \boldsymbol{\theta}_0, \ell_0 + th_\ell, \boldsymbol{m}_0)\right]\Big|_{t=0} = 0$, and $\frac{\partial}{\partial t}\mathbb{E}\left[\psi(\boldsymbol{W}_{jt}; \boldsymbol{\theta}_0, \ell_0, \boldsymbol{m}_0 +\right.$ 0 for regular perturbations h_ℓ, h_m, ensuring first-order insensitivity to nuisance estimation errors. Direct computation establishes orthogonality in each direction: for the ℓ direction, $\partial_t \mathbb{E}[\psi] = -\mathbb{E}[h_\ell(\boldsymbol{X})\,\mathbb{E}[\boldsymbol{S} - \boldsymbol{m}_0(\boldsymbol{X}) \mid \boldsymbol{X}]] = 0$; for the $\boldsymbol{m}$ direction, the cross term vanishes because $\mathbb{E}[\boldsymbol{S} - \boldsymbol{m}_0(\boldsymbol{X}) \mid \boldsymbol{X}] = \boldsymbol{0}$ and the residual term vanishes because $\mathbb{E}[\varepsilon \mid \boldsymbol{X}] = 0$. The $\sqrt{n}$-asymptotic theory in step (5) below requires $\hat{\ell}$ and $\hat{\boldsymbol{m}}$ to satisfy the high-level nuisance-rate condition, with products of nuisance errors small enough for the second-order remainder to be $o_p(n^{-1/2})$; a common sufficient condition is $o_p(n^{-1/4})$ convergence in appropriate L_2 norms, together with standard moment and nondegeneracy conditions (Chernozhukov et al., 2018, Assumption 3.2). We use random forests as flexible first-stage estimators, but do not claim that this implementation mechanically verifies those rates in the present finite sample. The DML estimates should therefore be read as robustness to nonlinear observables under the maintained high-level DML conditions. The nuisance parameters are estimated via a plug-in method that employs machine learning techniques. In our case, we employ an ensemble learning method that combines multiple decision trees to produce robust and flexible predictions for regression tasks. Each tree is trained on a bootstrap sample of the data, and predictions are aggregated to reduce variance and improve generalization.

Relative to Ordinary Least Squares (OLS), DML allows $\boldsymbol{X}_{jt}$ to enter through a general, unknown function $g_0(\cdot)$. This method removes first-order sensitivity of the target to nuisance estimation error via orthogonalization. We mitigate the problem of overfitting through sample splitting and out-of-fold prediction. Specifically, our estimator is based on a K-fold cross-fitting procedure following Chernozhukov et al. (2018):

1. **Split.** Randomly partition the sample $\{1, \ldots, n\}$ into K disjoint folds $I_1, ..., I_K$, each containing approximately n/K observations.

2. **Train nuisances off-fold.** For each fold k, fit $\hat{\ell}^{(-k)}(\boldsymbol{x})$ and $\hat{\boldsymbol{m}}^{(-k)}(\boldsymbol{x})$ on the complement sample I_k^c using plug-in machine learning methods.

3. **Predict on held-out fold.** For each $(j, t) \in I_k$, compute the cross-fit residuals using the

out-of-fold nuisance estimators:

$$\tilde{Y}_{jt} = \bar{Y}_{jt} - \hat{\ell}^{(-k)}(\boldsymbol{X}_{jt}), \quad \tilde{\boldsymbol{S}}_{jt} = \boldsymbol{S}_{jt} - \hat{\boldsymbol{m}}^{(-k)}(\boldsymbol{X}_{jt}),$$

where $\hat{\ell}^{(-k)}$ and $\hat{\boldsymbol{m}}^{(-k)}$ are machine learning estimates of ℓ_0 and $\boldsymbol{m}_0$ trained on the complement of fold k.

4. **Aggregate and estimate.** Stack residuals across all folds and compute $\hat{\boldsymbol{\theta}}$ using the following equation:

$$\hat{\boldsymbol{\theta}} = \left(\frac{1}{n} \sum_{j,t} \tilde{\boldsymbol{S}}_{jt} \tilde{\boldsymbol{S}}'_{jt} \right)^{-1} \left(\frac{1}{n} \sum_{j,t} \tilde{\boldsymbol{S}}_{jt} \tilde{Y}_{jt} \right).$$

5. **Variance.** The asymptotic variance of our estimators is evaluated using the sandwich form:

$$\widehat{\mathrm{Var}}(\hat{\boldsymbol{\theta}}) = \hat{J}^{-1} \hat{\Sigma} \hat{J}^{-1} / n,$$

where $\hat{J} = \frac{1}{n} \sum_{j,t} \tilde{\boldsymbol{S}}_{jt} \tilde{\boldsymbol{S}}'_{jt}$, and the estimated covariance matrix is $\hat{\Sigma} = \frac{1}{n} \sum_{j,t} \psi_{jt}(\hat{\boldsymbol{\theta}}) \psi_{jt}(\hat{\boldsymbol{\theta}})'$, where $\psi_{jt}(\hat{\boldsymbol{\theta}}) = \tilde{\boldsymbol{S}}_{jt}\big(\tilde{Y}_{jt} - \tilde{\boldsymbol{S}}'_{jt}\hat{\boldsymbol{\theta}}\big)$.

E Data clean-up and exclusion process

This appendix explains the steps taken to turn the original data into the net sample used in the baseline analysis. The descriptions below closely follow Table A3. The raw 2013–2014 sample consists of 36,576 request-level observations. Subsidy requests could be for one year, a fraction of a year, or multiple years. To mitigate the potential biases that this variability could induce, we discarded non-annual contract. Table A4 column r3 confirms that this exclusion step is inconsequential.

Table A3—Data clean-up steps.

Data clean-up step	Sample size
Original sample	36,576
Limit to annual contracts	13,370
Limit category: Data; Leased/Tariffed Facilities or Services	13,350
Drop grandfathered: P1 obs with missing price info	12,080
Removed requests that were submitted or withdrawn	12,038
Remove cases for which we could not recover price	11,798
Drop Alaska	11,308
Drop if speed is missing	11,290
Limit to valid speed units	11,245
Keep speed above 50Kbps	11,234
Limit to speeds between 1-100 Mbps	8,050
Drop subsidy = 0	7,780
One observation per HCP-year (TWFE aggregation)	4,231
Limit to HCPs that appear in both years	1,940

Notes: Counts are request-level through the speed and subsidy filters and become HCP-year-level after aggregation. Filters apply to every year, but counts shown reflect the 2013–2014 baseline pair.

While $\mathcal{P}_2/\mathcal{P}_2^c$ subsidizes only internet plans, $\mathcal{P}_1$ may also cover other categories such as equipment, construction, infrastructure, management, and maintenance. To avoid contamination, we limit the $\mathcal{P}_1$ sample to two categories of expense that refer to the cost of internet plans, "Data" or "Leased/Tariffed Facilities or Services".

Since the inception of the program in 1997, the FCC has reclassified rurality boundaries several times. Each reclassification causes two events. Some formerly urban HCPs would become rural and eligible for subsidies, whereas some formerly rural HCPs would become urban. HCPs in the latter group may be grandfathered to remain eligible. If the latter group chooses to remain in $\mathcal{P}_1$, they are no longer obligated to report a comparison of rural and urban prices because they are now located in an urban area. Given that in $\mathcal{P}_1$, the difference between rural and urban prices is key to recovering the subsidy amount, and it is unobservable for grandfathered HCPs, we discarded grandfathered observations in $\mathcal{P}_1$.

There are multiple values for request status in the data. Requests are initially "submitted". A request may be a "duplicate" of another request. HCPs may choose to "withdraw" a request. Otherwise, a request undergoes processing to be eventually "approved", "partially approved", or "denied". A "committed" status means that the request has been approved and the fund disbursement is done or underway. That said, we limited the sample to "committed" requests.

Next, we limit the sample to observations for which we were able to recover both subsidy and price. The recovery process is different for $\mathcal{P}_1$ vs $\mathcal{P}_2/\mathcal{P}_2^c$. Under $\mathcal{P}_1$, a subsidy request must mention the rural and urban price. We set the price equal to the rural price, and we set the subsidy equal to the difference between the rural and urban prices. Under $\mathcal{P}_2/\mathcal{P}_2^c$, the request specifies the subsidy amount. In this case, we take the subsidy amount as given, and we set the price equal to the subsidy divided by 0.65.

Next, we drop HCPs that are located in Alaska. Alaskan observations include many outliers featuring extremely high prices and very low speeds. This is an artifact of Alaska's unique demography and geography. Alaska has the 4th lowest population among US states, but it is the largest state by area (2.47 times the size of the second largest state, Texas).[11] While the average population density in the US is 37 people per square kilometer, the population density in Alaska is 0.50.[12] In per-capita terms, it is prohibitively costly to provide landline internet connection to such a sparsely populated state across vast geographies and rough terrain. So, expensive technologies such as satellite and Multiprotocol Label Switching (MPLS), that make up only 5% of users in other states, make up 81% of the users in Alaska. As a result, Alaskan internet is 20 times costlier than the national average and it ranks last in terms of speed. That said, following (Rabbani, 2024a,b), we exclude Alaska from our baseline analysis to ensure that the results are not influenced by the anomalies of Alaskan internet. Nevertheless, the exclusion of Alaska is inconsequential (Table A4, column r5).

Throughout this paper, we take download speed as the measure of internet speed, i.e., the measure of the quantity of consumption of internet. This assumption was made to overcome a data limitation. The data reports internet speed differently under each program. Under $\mathcal{P}_2/\mathcal{P}_2^c$, both download speed and upload speed are reported. Under $\mathcal{P}_1$, the data reports bandwidth as a single measure of speed. We believe this corresponds to download speed because it is usually the case that, for a given internet plan, download speed is greater than or equal to the upload speed.[13] We verified that this condition holds in 98.66% of the observable sample. That said, if all programs report download speed, and some programs ($\mathcal{P}_1$) do not report upload speed, then the only available measure of the quantity of consumption would be download speed.

However, recovering download speed was a separate challenge. The data reports bandwidth and download speed as a free-form user input that combines speed and speed unit. This enables non-standard inputs (reporting in gigabits per second, megabits per second, etc.) and erroneous data entry such as invalid speed units. To address this, we limited the sample to observations for which we could identify one of the following speed units: "GB", "Gbps", "KB", "Kbps", "MB", "Mbps", or "Gigabyte per second". Then we converted all speed units to the megabit per second (Mbps) equivalent. We discarded observations for which speed was missing or the speed unit was not recoverable. To ensure that this data exclusion step is not distorting the data, Table A4,

[11] List of U.S. states and territories by area, List of U.S. states and territories by population
[12] List of U.S. states and territories by population density
[13] Based on the Information Technology and Innovation Foundation

column (r2) limits the sample to observations that verbatim specified speed in "Mbps". The results confirm the baseline findings.

Since the program still subsidizes legacy technologies, very low speeds, such as dial-up connections, are commonly found in the data. On the other hand, some HCPs have established data centers with extremely high speeds. We limit the sample to speeds that are in 1–100Mbps. The robustness check in Table A4, column (r4) shows that the results hold if all speed levels are included.

Next, we drop observations for which the subsidy amount is zero, as these do not represent subsidized internet access. The next step is the process of aggregating request-level data to HCP-level. A given HCP is allowed to request as many lines of internet subsidy as they need. Many HCPs have tens or hundreds of concurrent lines of subsidy. The data reports year and HCP ID. Using these two measures, we aggregate all requests of the same HCP in the same year into one observation. That is, the HCP-level internet speed is the sum of internet speeds that share the same HCP ID and year. A similar approach defines HCP-level measures of price and subsidy.

After aggregating to the HCP-year level, the 2013–2014 sample contains 4,231 observations. But not all HCPs that appear in 2013 appear in 2014, and vice versa. To create a balanced panel, we limit the sample to HCPs that appear in both years, yielding a net sample of 1,940 observations.

F Robustness checks

This appendix discusses the robustness of the baseline findings to alterations to the sample and specification. Table A4 reports the results for nine variants, each modifying one aspect of the baseline. The columns are discussed in order below.

Table A4—Robustness checks: continuous treatment, 2014.

	Excluding medical schools (r1)	**Mbps speeds only** (r2)	**All contract durations** (r3)	**All speed ranges** (r4)	**Including Alaska** (r5)	**Rural health clinics** (r6.1)	**Non-profit hospitals** (r6.2)	**Community health centers** (r6.3)
Panel A: ln(price)								
$\tau_{1,2}$	−1.252	−1.821	−0.878	−1.205	−1.274	−0.692	−0.982	−4.693
	(0.086)	(0.066)	(0.062)	(0.071)	(0.078)	(0.134)	(0.100)	(0.439)
$\tau_{1,2^c}$	0.555	0.446	−0.223	0.497	0.548	0.518	0.735	−0.908
	(0.136)	(0.071)	(0.091)	(0.142)	(0.128)	(0.108)	(0.208)	(0.743)
$\tau_{1,2^c} - \tau_{1,2}$	1.806	2.267	0.655	1.701	1.822	1.210	1.717	3.785
	(0.147)	(0.092)	(0.098)	(0.149)	(0.139)	(0.153)	(0.208)	(0.923)
R^2	0.385	0.686	0.467	0.495	0.388	0.423	0.526	0.848
Panel B: ln(subsidy)								
$\tau_{1,2}$	−1.259	−1.778	−0.355	−1.279	−1.277	−0.800	−0.883	−5.183
	(0.103)	(0.083)	(0.041)	(0.084)	(0.094)	(0.164)	(0.129)	(0.477)
$\tau_{1,2^c}$	0.283	0.167	0.146	0.252	0.287	0.192	0.558	−1.237
	(0.164)	(0.090)	(0.060)	(0.169)	(0.154)	(0.132)	(0.268)	(0.807)
$\tau_{1,2^c} - \tau_{1,2}$	1.542	1.945	0.501	1.531	1.564	0.993	1.441	3.946
	(0.178)	(0.116)	(0.065)	(0.176)	(0.167)	(0.187)	(0.269)	(1.002)
R^2	0.310	0.564	0.457	0.464	0.313	0.273	0.435	0.850
Panel C: ln(HCP net cost)								
$\tau_{1,2}$	−0.720	−1.541	−1.019	−0.615	−0.752	−0.293	−0.777	−2.263
	(0.075)	(0.092)	(0.104)	(0.064)	(0.069)	(0.203)	(0.102)	(0.343)
$\tau_{1,2^c}$	1.739	1.674	0.078	1.631	1.699	1.868	1.634	0.586
	(0.119)	(0.100)	(0.151)	(0.127)	(0.113)	(0.163)	(0.213)	(0.580)
$\tau_{1,2^c} - \tau_{1,2}$	2.459	3.214	1.098	2.247	2.452	2.161	2.411	2.850
	(0.128)	(0.129)	(0.163)	(0.133)	(0.123)	(0.231)	(0.213)	(0.721)
R^2	0.433	0.589	0.406	0.506	0.424	0.496	0.494	0.710
N	1,934	1,432	4,882	2,506	2,238	530	804	192

Notes: Specification matches the "Cont." columns of Table 4. Each column applies one sample restriction to the 2014 baseline. The number of observations is identical across the three panels. Standard errors in parentheses.

Excluding medical schools (r1). The baseline sample includes a small number of post-secondary educational institutions (medical schools). Because these entities differ from typical HCPs in size and internet needs, column (r1) excludes them. The results are virtually identical to the baseline.

Mbps speeds only (r2). As discussed in Section E, converting free-form speed entries to a common unit required assumptions about non-standard speed units (e.g., GB, MB, KB). If certain errors are clustered in one program, they could introduce bias. Column (r2) restricts the sample to observations that verbatim reported speed in Mbps, eliminating any unit-conversion assumptions. The results support the baseline findings.

All contract durations (r3). The baseline excludes non-annual subsidy contracts. For example, if an HCP began receiving subsidies mid-year, the baseline discards the partially funded year. Column (r3) retains all contract durations. The sample size substantially rises, and the results remain qualitatively unchanged.

All speed ranges (r4). The baseline limits speeds to 1–100Mbps to mitigate the effect of outliers at both ends. Legacy technologies such as dial-up sit at the low end, and data-center-grade connections at the high end. Column (r4) removes the speed restriction entirely. If anything, the results become larger and more significant, strongly supporting the baseline findings.

Including Alaska (r5). We excluded Alaska from the baseline for the reasons explained in Section E. Column (r5) retains Alaskan observations. The results remain qualitatively intact.

Single HCP types (r6.1–r6.3). Different HCP types may have different internet needs, and a concentration of specific types in one program could bias the estimates. To investigate, we estimate the effects within three HCP types that have sufficient sample sizes across programs: rural health clinics (r6.1), non-profit hospitals (r6.2), and community health centers (r6.3). The results corroborate the baseline findings.

Ethernet only (r7). The baseline keeps all internet service types, listed in Table A7. Some technologies may be inherently more expensive for reasons such as reliability or responsiveness, and clustering of expensive technologies in one program could induce bias. Column (r7) limits the data to Ethernet, the most common service type found across all three programs. The much smaller sample reduces statistical power, and several coefficients cannot be estimated. Yet, to the extent that estimates could be produced, the results confirm the baseline findings.

F.1 Pre-period speed control for the pooled OLS

The pooled OLS in Section 5.2 (Table 3) conditions on observation-level ln(speed), which is contemporaneous with the program-choice outcome. Under the model in Section 3, switching from $\mathcal{P}_1$ to $\mathcal{P}_2$ can raise the quantity consumed (Proposition **P.2**, Part ii), so 2014 speed may be partly mechanism-driven. To rule out post-treatment conditioning, Table A5 re-estimates the same pooled OLS using $\ln(\text{speed})_{2013}$ as a time-invariant covariate. We define $\ln(\text{speed})_{2013}$ as the log of each HCP's *total* 2013 bandwidth, summed across all 2013 subsidy requests of the HCP (any group) and then logged. Because the variable is fixed within HCP across the 2013–2014 panel, this specification rules out any post-treatment conditioning channel operating through contemporaneous speed.

All 970 baseline HCPs have non-missing 2013 bandwidth, so the regression sample is unchanged ($N = 4{,}234$). The cross-subsidization premium $\beta_3 - \beta_2$ remains highly significant across all three outcomes ($p < 0.001$), with magnitudes about 26–30% smaller than under the contemporaneous control. The premium falls from 1.58 to 1.11 for ln(price), from 1.34 to 0.94 for ln(subsidy), and from 1.87 to 1.38 for ln(HCP net cost). The reduction reflects a modest selection-on-baseline-bandwidth channel. $\mathcal{P}_2^c$ participants tend to be larger-bandwidth users in the pre-period, and absorbing this composition through the time-invariant control shrinks the apparent premium. The qualitative cross-subsidization conclusion is unchanged.

Table A5—Pooled OLS with pre-period ln(speed) control, 2013–2014.

	ln(price)	ln(subsidy)	ln(HCP net cost)
β_1	7.6549	7.5039	5.7128
	(0.1384)	(0.1769)	(0.2265)
β_2	6.6608	6.4485	5.4087
	(0.1309)	(0.1673)	(0.2141)
β_3	7.7726	7.3896	6.7883
	(0.1717)	(0.2194)	(0.2809)
$\beta_3 - \beta_2$	1.1118	0.9411	1.3796
	(0.1107)	(0.1415)	(0.1811)
ln(speed)$_{2013}$	−0.0172	−0.0433	0.0292
	(0.0084)	(0.0108)	(0.0138)
N	4,234	4,234	4,234
Adj. R^2	0.9941	0.9895	0.9783

Notes: Robustness variant of Table 3 with the time-varying observation-level ln(speed) control replaced by ln(speed)$_{2013}$, defined as the log of each HCP's total 2013 bandwidth summed across requests. Because this regressor is fixed within HCP, the specification rules out post-treatment conditioning through contemporaneous speed. Other specification details follow Table 3. Sample: 970 baseline HCPs with non-missing 2013 bandwidth and positive HCP net cost. Standard errors in parentheses.

F.2 Speed-control sensitivity for the headline TWFE/DML table

The headline regression in Table 4 conditions on observation-level ln(speed), which is contemporaneous with the program-choice outcome. As noted in Section F.1, Proposition **P.2**, Part (ii), implies that 2014 speed may be partly mechanism-driven. Under HCP fixed effects, a time-invariant pre-period speed measure is perfectly collinear with the fixed effects and cannot enter directly, so the cleanest analog of the pooled-OLS robustness is to drop ln(speed) entirely. The size control ln(requests) is retained.

Table A6 reports the headline TWFE-Continuous, TWFE-Binary, and DML estimates without the speed control. The substitution recovers the *total* mechanism effect on prices, subsidies, and HCP net cost, including any quantity-channel response, rather than the partial effect at constant within-HCP speed change. The $\tau_{1,2}$ coefficients shrink substantially across all three outcomes (40%–60% in absolute value). The $\mathcal{P}_1 \rightarrow \mathcal{P}_2$ effect on ln(price) falls from -1.249^{***} to -0.734^{***} in Panel A, on ln(subsidy) from -1.262^{***} to -0.728^{***} in Panel B, and on ln(HCP net cost) from -0.704^{***} to -0.265^{***} in Panel C. The cross-subsidization premium $\tau_{1,2^c} - \tau_{1,2}$ shrinks more modestly (13%–26%) and remains highly significant ($p < 0.001$) in every outcome and every panel. The qualitative conclusions of Section 5.3 (cost-effectiveness of $\mathcal{P}_2$ over $\mathcal{P}_1$, and a cross-subsidization premium under $\mathcal{P}_2^c$) survive the speed-control sensitivity, though the magnitudes are more conservative.

The magnitude reduction has a clean economic interpretation. The original specification measures the per-unit-speed price effect. The no-ln(speed) specification measures the total dollar effect, including the upward shift in quantity demanded. Switching from $\mathcal{P}_1$ to $\mathcal{P}_2$ lowers per-unit price by approximately $1 - e^{-1.249} \approx 71\%$ but also allows the HCP to buy more bandwidth, so the total spending reduction is closer to $1 - e^{-0.734} \approx 52\%$. Both magnitudes are defensible. The policy-relevant interpretation depends on whether the welfare comparison is held at constant quantity (the per-unit measure) or at the equilibrium quantity (the total measure).

Table A6—Combined regression results without ln(speed), 2014.

	Panel A: ln(**price**)			Panel B: ln(**subsidy**)			Panel C: ln(**HCP net cost**)		
	Cont.	Binary	DML	Cont.	Binary	DML	Cont.	Binary	DML
$\tau_{1,2}$	−0.734	−0.752	−0.988	−0.728	−0.736	−0.905	−0.265	−0.261	−0.571
	(0.060)	(0.065)	(0.156)	(0.072)	(0.078)	(0.165)	(0.052)	(0.056)	(0.154)
$\tau_{1,2^c}$	0.684	0.745	0.750	0.417	0.464	0.537	1.850	1.981	1.677
	(0.140)	(0.163)	(0.312)	(0.167)	(0.195)	(0.318)	(0.122)	(0.138)	(0.306)
$\tau_{1,2^c} - \tau_{1,2}$	1.418	1.496	1.738	1.145	1.200	1.442	2.115	2.243	2.248
	(0.144)	(0.168)	(0.382)	(0.172)	(0.201)	(0.381)	(0.126)	(0.143)	(0.422)
N	1,940	1,670	1,940	1,940	1,670	1,940	1,940	1,670	1,940
R^2	0.343	0.345		0.275	0.271		0.392	0.411	

Notes: Robustness variant of Table 4 that drops ln(speed) from the covariate vector. Under HCP fixed effects, a pre-period (time-invariant) speed control is perfectly collinear with the HCP fixed effects, so dropping the time-varying speed control is the closest analog of Table A5's pre-period-speed robustness. The size control ln(requests) is retained. Other specification details follow Table 4. Standard errors in parentheses.

G Supporting tables and figures

Table A7—Summary statistics for the POLS model.

Variable	Obs	Mean	SE	Median	Min	Max	
$\mathcal{P}_1$	3,205						
$\mathcal{P}_2$	973						
$\mathcal{P}_2^c$	56						
Speed (Mbps)	4,234	13.14	0.350	1.54	1.50	100	
Price ($)	4,234	14,044	314.70	8,328	60.00	270,024	
Subsidy ($)	4,234	10,706	266.38	5,715	39.00	242,974	
HCP net cost ($)	4,234	3,338	76.45	2,106	0.00	87,462	
HCP type	**Obs**						
Non-profit hospital	2,352	Local health department	266	Emergency room	17		
Rural health clinic	808	Mental health center	265	Medical school	6		
Community health center	288	Not available	232				
Service type	**Obs**						
T1	2,084	MPLS	390	T3	75	Fiber	31
Ethernet	504	Internet	270	DSL	65	VPN	20
ISDN	472	PRI	246	Voice_Grade	36	Other categories	41
State	**Obs**						
GA	585	MN	106	NY	50	PA	20
WI	507	OH	105	NM	49	SC	16
TX	272	IA	100	OR	49	UT	13
CA	261	MO	94	LA	43	NH	11
VA	253	MI	72	MT	36	ND	10
NE	222	KS	67	OK	33	WV	5
KY	216	SD	62	WY	30	MA	4
MS	180	AL	61	CO	28	FL	2
AR	171	TN	57	NV	24		
IL	152	NC	53	WA	21		
AZ	122	IN	52	ID	20		

Notes: Sample of the 4,234 HCP-year observations from the 970 baseline HCPs over 2013–2014 used in Table 3. Categorical breakdowns report counts.

Table A8—Summary statistics of consortium members.

Variable	Obs	Mean	SE	Median	Min	Max
Mean price for eligible HCPs	628	11,340.5	262.6	10,312.1	321.0	63,216.5
Mean speed for eligible HCPs	628	1,397.0	55.8	996.4	1.5	8,192.0
Mean number of bidders	628	0.849	0.055	0.140	0.000	9.938
Ineligible HCPs' share by speed	628	0.196	0.009	0.112	0.000	1.000
Ineligible HCPs' share by count	628	0.217	0.007	0.162	0.010	0.944
Consortium's total speed	628	77,133.0	7,007.6	25,274.9	12.4	1,801,124.7

Notes: The unit of observation is consortium-year (628 consortium-year cells pooled across 2014–2021). Each row reports a per-consortium aggregate. Consortium members may be eligible (subsidized via $\mathcal{P}_2^c$) or ineligible. Speeds in Mbps, prices in dollars.

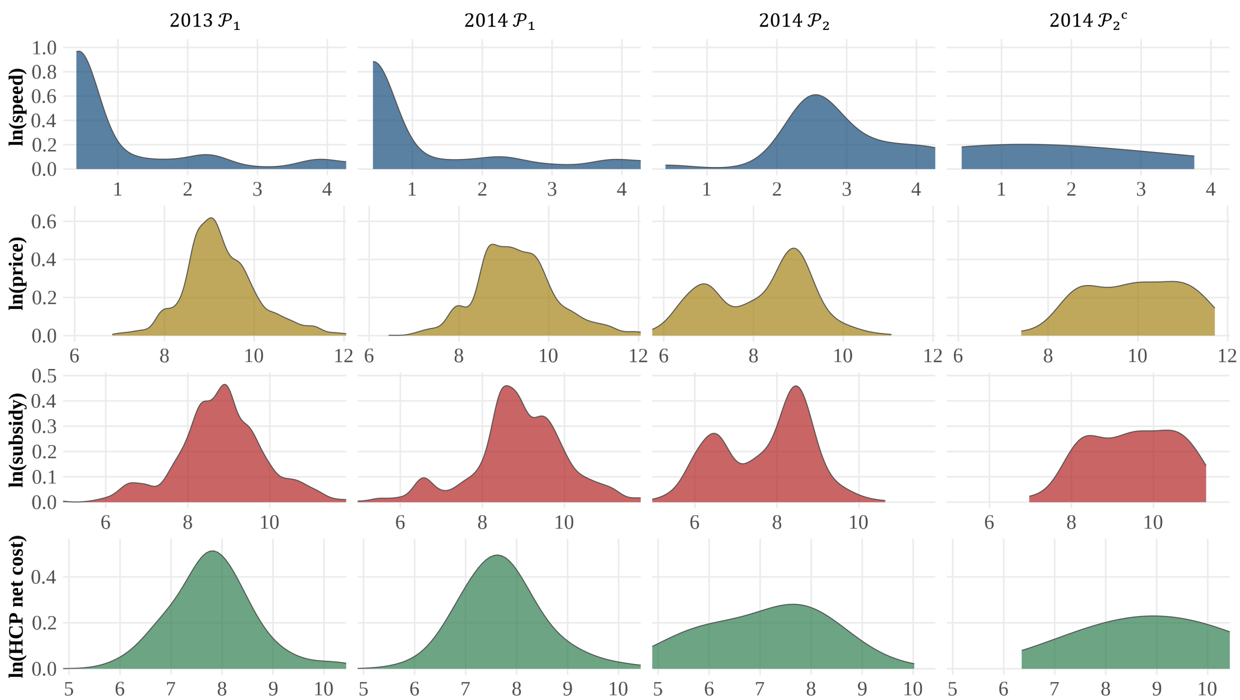

Figure A2: Kernel density of request-level variables by program, 2013–2014.

Notes: Request-level baseline sample (970 HCPs in 2013–2014). Rows: ln(speed), ln(price), ln(subsidy), ln(HCP net cost). Within each row, axes are held constant across columns to ease cross-program comparison.

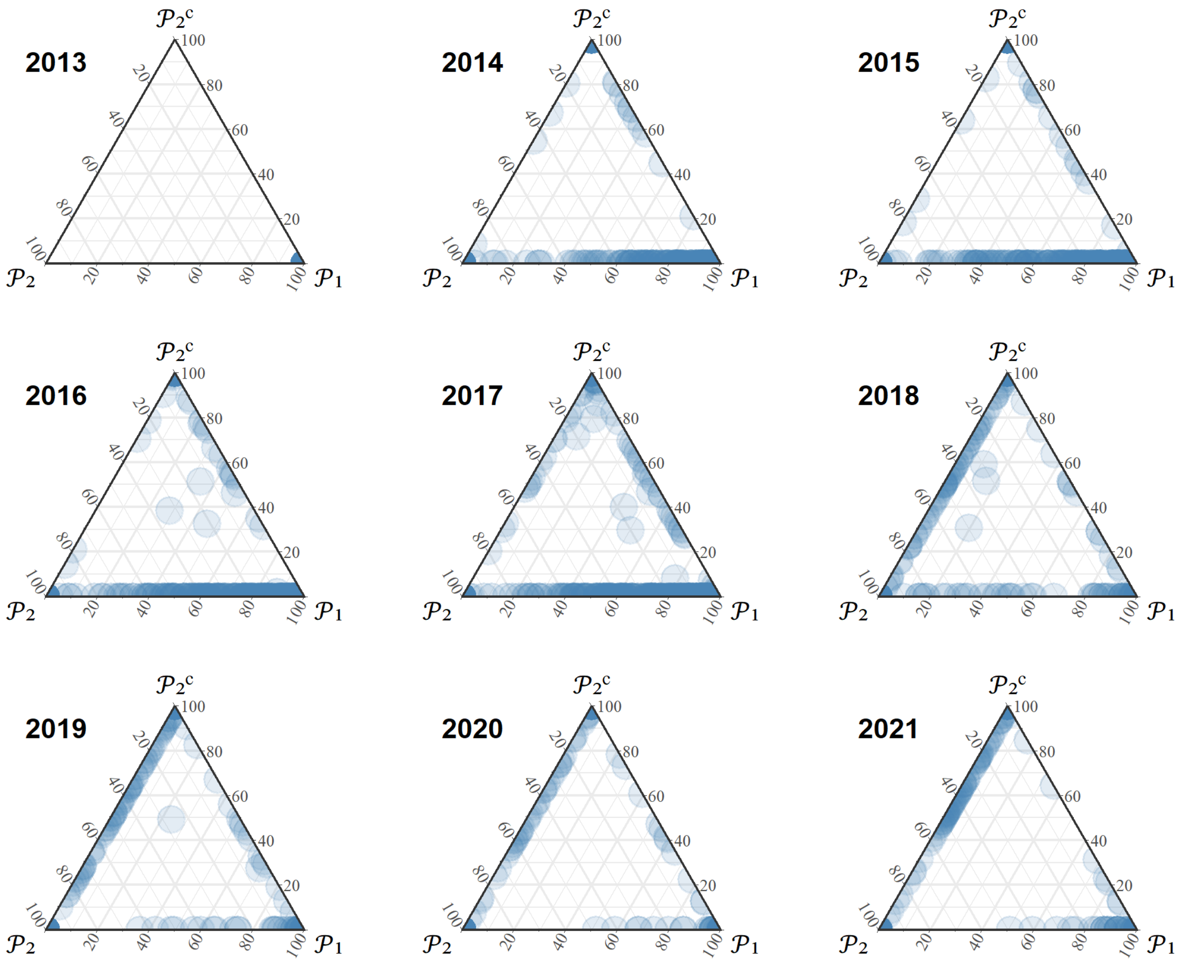

Figure A3: Program-share composition of HCPs, 2013–2021.

Notes: Each point is one HCP-year. The three axes give the fraction of the HCP's subsidized requests under $\mathcal{P}_1$, $\mathcal{P}_2$, and $\mathcal{P}_2^c$, summing to one. Proximity to a vertex indicates concentration under that program. Panels span 2013–2021, matching the aggregates in Figure 3.

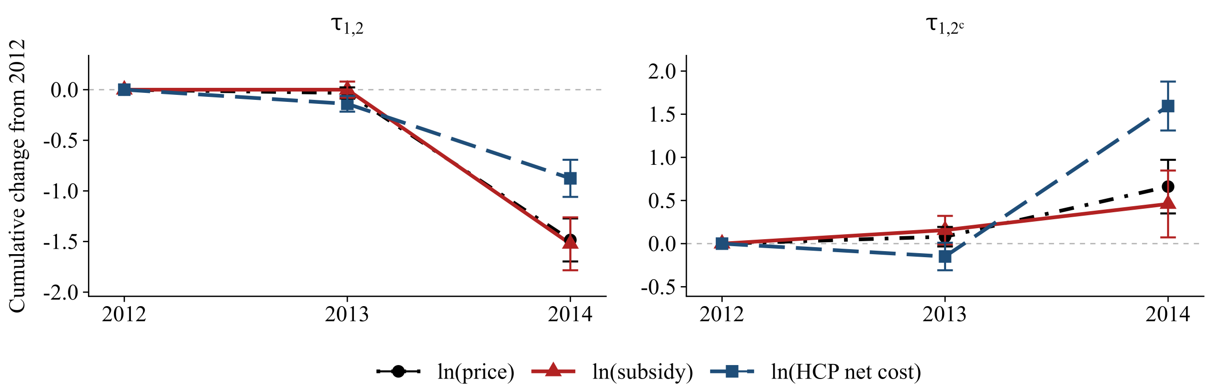

Figure A4: Three-period event-study coefficients, 2012–2014.

Notes: Visualization of Table 5. Lines trace cumulative log-changes from 2012 on the balanced panel of 642 HCPs. The 2013 values are placebo coefficients (should be near zero under parallel trends); the 2014 values add the headline treatment effects. Left panel: $\tau_{1,2}$ ($\mathcal{P}_1 \to \mathcal{P}_2$); right panel: $\tau_{1,2^c}$ ($\mathcal{P}_1 \to \mathcal{P}_2^c$). Whiskers are 95% confidence intervals. Specification details follow Table 5.

Table A9—TWFE continuous regression results by year pair.

	2014	2015	2016	2017	2018	2019	2020	2021
Panel A: ln(**price**)								
$\tau_{1,2}$	−1.249	−1.275	−1.044	−0.705	−0.499	−0.869	−0.949	−0.595
	(0.085)	(0.086)	(0.101)	(0.083)	(0.062)	(0.133)	(0.179)	(0.170)
$\tau_{1,2^c}$	0.556	0.008	−0.896	0.514	−0.328	−0.344	−0.873	−0.594
	(0.136)	(0.143)	(0.295)	(0.162)	(0.073)	(0.311)	(0.214)	(0.260)
$\tau_{1,2^c} - \tau_{1,2}$	1.804	1.284	0.149	1.219	0.171	0.525	0.076	0.001
	(0.147)	(0.162)	(0.305)	(0.169)	(0.069)	(0.331)	(0.267)	(0.292)
N	1,940	3,658	3,950	4,388	5,832	7,004	7,290	7,812
R^2	0.387	0.410	0.409	0.366	0.362	0.271	0.270	0.278
Panel B: ln(**subsidy**)								
$\tau_{1,2}$	−1.262	−1.398	−1.118	−0.734	−0.563	−0.951	−0.735	−0.757
	(0.102)	(0.093)	(0.105)	(0.086)	(0.063)	(0.134)	(0.181)	(0.170)
$\tau_{1,2^c}$	0.284	0.201	−1.090	0.474	−0.507	−0.623	−0.850	−0.693
	(0.164)	(0.153)	(0.306)	(0.168)	(0.074)	(0.314)	(0.217)	(0.261)
$\tau_{1,2^c} - \tau_{1,2}$	1.546	1.599	0.029	1.208	0.056	0.328	−0.115	0.065
	(0.177)	(0.174)	(0.317)	(0.175)	(0.070)	(0.334)	(0.271)	(0.293)
N	1,940	3,658	3,950	4,388	5,832	7,004	7,290	7,812
R^2	0.312	0.402	0.403	0.356	0.387	0.269	0.266	0.279
Panel C: ln(**HCP net cost**)								
$\tau_{1,2}$	−0.704	−0.390	−0.578	−0.285	−0.084	−0.524	−0.424	0.294
	(0.074)	(0.088)	(0.098)	(0.091)	(0.071)	(0.138)	(0.180)	(0.170)
$\tau_{1,2^c}$	1.740	0.100	0.263	0.848	0.247	0.958	−0.370	0.053
	(0.119)	(0.145)	(0.288)	(0.177)	(0.082)	(0.324)	(0.216)	(0.261)
$\tau_{1,2^c} - \tau_{1,2}$	2.445	0.490	0.841	1.133	0.331	1.482	0.054	−0.241
	(0.128)	(0.165)	(0.298)	(0.185)	(0.078)	(0.345)	(0.270)	(0.292)
N	1,940	3,658	3,950	4,388	5,832	7,004	7,290	7,812
R^2	0.431	0.321	0.375	0.313	0.305	0.262	0.255	0.273

Notes: Each column reports continuous-share Two-Way Fixed Effects (TWFE) estimates for year pair $(t-1, t)$, restricting to HCPs that appear in both years. Specification matches the "Cont." columns of Table 4. Standard errors in parentheses.

Table A10—Goodness of fit: price–speed functional forms, 2014.

Rank	Model	k	$\mathcal{P}_1$ Adj. R^2	$\mathcal{P}_1$ RMSE	$\mathcal{P}_2$ Adj. R^2	$\mathcal{P}_2$ RMSE	$\mathcal{P}_2^c$ Adj. R^2	$\mathcal{P}_2^c$ RMSE
Panel A: HCP level (N: $\mathcal{P}_1$=643, $\mathcal{P}_2$=419, $\mathcal{P}_2^c$=43)								
1	$P = a + bS + cS^2$	3	0.273	$22,930	0.009	$5,499	0.424	$21,209
2	$P = a + b \cdot \ln(S)$	2	0.242	$23,434	0.008	$5,508	0.470	$20,592
3	$P = a + b \cdot \sqrt{S}$	2	0.201	$24,058	0.007	$5,510	0.466	$20,669
4	$\ln(P) = a + b\ln(S) + c[\ln(S)]^2$	3	0.168	$24,538	−0.121	$5,848	0.391	$21,814
5	$\ln(P) = a + b \cdot \ln(S)$	2	0.160	$24,680	−0.121	$5,856	0.424	$21,472
6	$P = F + c \cdot S$	2	0.138	$24,998	0.005	$5,516	0.430	$21,355
7	$\ln(P) = a + b \cdot \sqrt{S}$	2	0.098	$25,571	−0.120	$5,852	0.362	$22,607
8	$\ln(P) = a + b \cdot S$	2	0.033	$26,470	−0.123	$5,860	0.318	$23,368
Panel B: Request level (N: $\mathcal{P}_1$=1,305, $\mathcal{P}_2$=973, $\mathcal{P}_2^c$=56)								
1	$P = a + bS + cS^2$	3	0.318	$19,516	0.019	$6,457	0.558	$20,644
2	$P = a + b \cdot \ln(S)$	2	0.271	$20,180	−0.001	$6,524	0.576	$20,419
3	$P = a + b \cdot \sqrt{S}$	2	0.218	$20,903	0.000	$6,522	0.544	$21,170
4	$\ln(P) = a + b\ln(S) + c[\ln(S)]^2$	3	0.210	$20,999	−0.135	$6,943	0.521	$21,496
5	$\ln(P) = a + b \cdot \ln(S)$	2	0.178	$21,425	−0.147	$6,985	0.495	$22,290
6	$P = F + c \cdot S$	2	0.150	$21,784	−0.001	$6,525	0.488	$22,426
7	$\ln(P) = a + b \cdot \sqrt{S}$	2	0.105	$22,362	−0.143	$6,973	0.418	$23,912
8	$\ln(P) = a + b \cdot S$	2	0.043	$23,124	−0.145	$6,978	0.389	$24,508

Notes: Eight functional forms fitted per program on 2014 baseline data (970 HCPs). Panel A: HCP–program means; Panel B: request-level. P is annual price ($), S is download speed (Mbps), k counts estimated parameters. Adjusted R^2 and RMSE are computed on the price scale for all models; log-space predictions are back-transformed via exponentiation, so their price-scale R^2 can be negative (the back-transformed fit then predicts worse than the sample mean). Rows are ranked by Panel A's $\mathcal{P}_1$ adjusted R^2.

Table A11—Box-Cox test for the price–speed functional form, 2014.

Program	N	$\hat{\lambda}$	$H_0: \lambda = 0$ LR stat.	$H_0: \lambda = 0$ p-value	$H_0: \lambda = 1$ LR stat.	$H_0: \lambda = 1$ p-value
Panel A: HCP level						
$\mathcal{P}_1$	643	−0.170	22.60	<0.001	1,217.22	<0.001
$\mathcal{P}_2$	419	0.200	21.32	<0.001	321.35	<0.001
$\mathcal{P}_2^c$	43	0.190	1.59	0.208	29.08	<0.001
Panel B: Request level						
$\mathcal{P}_1$	1,305	0.070	12.43	<0.001	2,323.49	<0.001
$\mathcal{P}_2$	973	0.140	28.02	<0.001	1,054.35	<0.001
$\mathcal{P}_2^c$	56	0.240	4.22	0.040	39.52	<0.001

Notes: Box-Cox test for the dependent-variable transformation in the price–speed regression $P^{(\lambda)} = a + b\ln(S) + \varepsilon$, where $P^{(\lambda)} = (P^{\lambda} - 1)/\lambda$ (Box and Cox, 1964). The transformation nests log-log ($\lambda = 0$) and lin-log ($\lambda = 1$) as special cases. $\hat{\lambda}$ is the maximum likelihood estimate; the LR statistic is $2(\ell(\hat{\lambda}) - \ell(\lambda_0))$, distributed $\chi^2(1)$ under the null. Sample: 2014 baseline data (970 HCPs), matching Table A10. Panel A: HCP–program means; Panel B: request-level.

Table A12—Combined regression results (levels), 2014.

		Price	Subsidy	HCP net cost
$\tau_{1,2}$	Coeff.	−28.855	−22.581	−6.274
	SE	(3.795)	(3.433)	(0.829)
	Semi-elasticity	−0.933	−0.956	−0.861
$\tau_{1,2^c}$	Coeff.	20.171	7.741	12.431
	SE	(6.088)	(5.508)	(1.329)
	Semi-elasticity	0.652	0.328	1.706
$\tau_{1,2^c} - \tau_{1,2}$	Coeff.	49.027	30.322	18.705
	SE	(6.572)	(5.945)	(1.435)
	Semi-elasticity	1.586	1.283	2.568
N		1,940	1,940	1,940
R^2		0.192	0.135	0.290

Notes: Specification matches the "Cont." columns of Table 4, with dependent variables in dollar levels rather than logs. Coefficients and standard errors are in thousands of dollars. Semi-elasticity is $\hat{\beta}/\bar{y}$, the coefficient divided by the sample mean of the dependent variable.

Table A13—Semi-elasticity comparison: log-linear vs. linear models.

	Price		Subsidy		HCP net cost	
	Log	Level/$\bar{y}$	Log	Level/$\bar{y}$	Log	Level/$\bar{y}$
$\tau_{1,2}$	−1.249	−0.933	−1.262	−0.956	−0.704	−0.861
$\tau_{1,2^c}$	0.556	0.652	0.284	0.328	1.740	1.706
$\tau_{1,2^c} - \tau_{1,2}$	1.804	1.586	1.546	1.283	2.445	2.568
$\bar{y}$		30,915		23,630		7,285

Notes: Two ways of computing the implied semi-elasticity. Both columns use the continuous-share Two-Way Fixed Effects (TWFE) specification. "Log" is the coefficient from the log-linear model (Table 4), which equals the semi-elasticity $\partial \ln y/\partial x$ directly. "Level/$\bar{y}$" divides the coefficient from the linear model (Table A12) by the sample mean $\bar{y}$. $\bar{y}$ is the arithmetic mean across both pre- and post-treatment periods in 2014.

H Influence diagnostics

This appendix applies Cook's Distance to the baseline 2013–2014 Two-Way Fixed Effects (TWFE) specification to assess whether the estimates are driven by a small number of influential observations. For a fixed-effects panel model, Cook's Distance is computed on the within-transformed (HCP-demeaned) data. Let D_i denote the Cook's Distance for observation i. Following the standard threshold, observations with $D_i > 4/N$ are flagged as influential. Because influence depends on the outcome variable, we compute Cook's Distance separately for each outcome (ln price, ln subsidy, ln HCP net cost) and take the union of flagged observations across all three equations. To maintain the balanced panel structure, if either panel row (pre- or post-treatment) of an HCP is flagged, we drop the entire HCP.

Table A14 reports the baseline and trimmed estimates. Of the 970 baseline HCPs, 143 (14.7%) have at least one flagged observation. The flagging rate varies sharply by switching behavior. Only 21 of 511 stayers (4.1%) are flagged, compared to 95 of 416 $\mathcal{P}_1 \to \mathcal{P}_2$ switchers (22.8%), 26 of 40 $\mathcal{P}_1 \to \mathcal{P}_2^c$ switchers (65.0%), and 1 of 3 mixed switchers (33.3%). This pattern is expected. Switchers experience the largest within-HCP changes in treatment status and thus in prices, making them high-leverage observations by construction.

Table A14—Influence diagnostics: baseline vs. trimmed estimates.

	ln(**price**)		ln(**subsidy**)		ln(**HCP net cost**)	
	Baseline	Trimmed	Baseline	Trimmed	Baseline	Trimmed
$\tau_{1,2}$	−1.249	−0.830	−1.262	−0.857	−0.704	−0.454
	(0.085)	(0.050)	(0.102)	(0.069)	(0.074)	(0.067)
Δ%		+33.5%		+32.1%		+35.5%
$\tau_{1,2^c}$	0.556	0.503	0.284	0.176	1.740	1.723
	(0.136)	(0.103)	(0.164)	(0.142)	(0.119)	(0.137)
Δ%		−9.6%		−38.1%		−1.0%
R^2	0.387	0.597	0.312	0.474	0.431	0.481
N	1,940	1,654	1,940	1,654	1,940	1,654

Notes: Specification matches the "Cont." columns of Table 4. Cook's Distance is computed on the within-transformed (HCP-demeaned) model for each outcome; observations with $D > 4/N$ are flagged as influential. "Trimmed" columns drop every HCP with any flagged panel-row in any outcome equation. Δ% is the percentage change in the trimmed coefficient relative to the baseline. Standard errors in parentheses.

After dropping the 143 flagged HCPs, $\tau_{1,2}$ decreases in magnitude by roughly one-third across all outcomes, but remains significant. The trimmed estimate of −0.830 for ln(price) implies that switching to $\mathcal{P}_2$ reduces prices by approximately 56% (versus 71% at baseline). The $\tau_{1,2^c}$ coefficient is more stable. For ln(HCP net cost), the trimmed estimate (1.723) is within 1% of the baseline (1.740). This is striking given that 65% of $\mathcal{P}_2^c$ HCPs were flagged. The within R^2 rises after trimming, confirming that the flagged observations were outliers in the residual space rather than in the treatment-effect space. In sum, the direction, significance, and economic interpretation of all treatment effects survive the removal of influential observations.

I Common-support restriction on speed

The GAM price–speed curves and kernel density plots (Figure A2) show that $\mathcal{P}_2^c$ participants operate in a narrower speed range than those in $\mathcal{P}_1$ or $\mathcal{P}_2$. If the baseline treatment effects are driven by price variation at extreme speeds where $\mathcal{P}_2^c$ has no representation, the comparison across programs may not be on common support. This appendix addresses this concern by restricting all programs to the speed domain of $\mathcal{P}_2^c$. We identify the range of pre-treatment (2013) download speeds among HCPs that participate in $\mathcal{P}_2^c$ as [1.5, 139.6] Mbps. We then drop all HCPs whose 2013 speed falls outside this range, regardless of their program assignment. This removes 21 of 970 baseline HCPs (2.2%), all from $\mathcal{P}_1$ and $\mathcal{P}_2$, leaving a sample of 949 HCPs observed on common support. Table A15 compares the baseline and restricted estimates. The coefficients are virtually unchanged and retain their statistical significance.

Table A15—Common-support restriction.

	ln(**price**)		ln(**subsidy**)		ln(**HCP net cost**)	
	Baseline	Restricted	Baseline	Restricted	Baseline	Restricted
$\tau_{1,2}$	−1.249	−1.224	−1.262	−1.250	−0.704	−0.673
	(0.085)	(0.089)	(0.102)	(0.107)	(0.074)	(0.076)
Δ%		+2.0%		+0.9%		+4.5%
$\tau_{1,2^c}$	0.556	0.556	0.284	0.288	1.740	1.727
	(0.136)	(0.137)	(0.164)	(0.166)	(0.119)	(0.117)
Δ%		+0.1%		+1.6%		−0.8%
R^2	0.387	0.366	0.312	0.298	0.431	0.415
N	1,940	1,898	1,940	1,898	1,940	1,898

Notes: Specification matches the "Cont." columns of Table 4. "Restricted" columns limit the sample to HCPs whose pre-treatment (2013) speed falls within the $\mathcal{P}_2^c$ range $[1.5, 139.6]$ Mbps. Δ% is the percentage change in the restricted coefficient relative to the baseline. Standard errors in parentheses.

J Coefficient stability: Oster bounds

This appendix applies the coefficient stability framework of Oster (2019) to assess whether unobserved time-varying confounders could explain away the estimated treatment effects. Consider two versions of the baseline TWFE model. The *short* model includes only the treatment structure and HCP fixed effects but excludes time-varying controls. The *long* model adds log speed and log number of requests. Let $\tilde{\beta}$ and $\hat{\beta}$ denote the treatment coefficient from the short and long models, respectively, and let $\tilde{R}^2$ and R^2 denote the corresponding within R^2 values. Under proportional selection, the degree to which selection on unobservables would need to exceed selection on observables to drive the treatment effect to zero is:

$$\delta = \frac{\hat{\beta} \cdot (R^2 - \tilde{R}^2)}{(\tilde{\beta} - \hat{\beta}) \cdot (R^2_{\max} - R^2)}$$

where $R^2_{\max} = \min(1.3 \times R^2, 1)$ following Oster (2019). The bias-adjusted coefficient assuming equal proportional selection ($\delta = 1$) is:

$$\beta^*(\delta = 1) = \hat{\beta} - (\tilde{\beta} - \hat{\beta}) \cdot \frac{R^2_{\max} - R^2}{R^2 - \tilde{R}^2}$$

Table A16 reports the results ($\tau_{2,2^c}$ is omitted because $\mathcal{P}_2$ was introduced in 2014, so no HCP could have been on $\mathcal{P}_2$ in 2013). δ is negative for every coefficient and outcome. For $\tau_{1,2}$, the uncontrolled $\tilde{\beta}$ is near zero or positive, but the controlled $\hat{\beta}$ is large and negative, indicating that adding controls *reveals* the treatment effect rather than inflating it. For unobservables to drive $\hat{\beta}$ to zero, they would have to work in the opposite direction of the observables, which is implausible under standard assumptions (Altonji, Elder, and Taber, 2005). $\tau_{1,2^c}$ is highly stable. δ ranges from -54 to -435 across outcomes, and $\beta^*(\delta = 1)$ barely differs from $\hat{\beta}$. These results complement the logit analysis in Table 6, which shows that observables contribute virtually no explanatory power to the switching decision beyond the mechanical cost ratio (pseudo R^2 from 0.031 to 0.037).

Table A16—Coefficient stability: Oster (2019) bounds.

	Panel A: ln(**price**)		**Panel B:** ln(**subsidy**)		**Panel C:** ln(**HCP net cost**)	
	$\tau_{1,2}$	$\tau_{1,2^c}$	$\tau_{1,2}$	$\tau_{1,2^c}$	$\tau_{1,2}$	$\tau_{1,2^c}$
$\tilde{\beta}$ (uncontrolled)	−0.261	0.543	−0.226	0.267	0.132	1.731
$\hat{\beta}$ (controlled)	−1.249	0.556	−1.262	0.284	−0.704	1.740
	(0.085)	(0.136)	(0.102)	(0.164)	(0.074)	(0.119)
$\tilde{R}^2$	0.043	0.043	0.015	0.015	0.135	0.135
R^2	0.387	0.387	0.312	0.312	0.431	0.431
$R^2_{\max}$	0.502	0.502	0.405	0.405	0.561	0.561
δ	−3.75	−126.81	−3.86	−53.67	−1.93	−434.58
$\beta(\delta=1)$	−1.582	0.560	−1.589	0.289	−1.070	1.744
N	1,940	1,940	1,940	1,940	1,940	1,940

Notes: Specification matches the "Cont." columns of Table 4. $\tilde{\beta}$ is the short-model coefficient (no time-varying controls); $\hat{\beta}$ is the long-model coefficient (adds ln(speed) and ln(requests)). $R^2_{\max} = \min(1.3 \times R^2, 1)$. δ is the proportional selection ratio needed to drive the effect to zero, and $\beta(\delta=1)$ is the bias-adjusted coefficient under equal selection on observables and unobservables. $\tau_{2,2^c}$ is omitted because $\mathcal{P}_2$ did not exist in 2013. Standard errors of $\hat{\beta}$ in parentheses.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1):151–184. URL https://doi.org/10.1086/426036.

Box, G. E. P. and D. R. Cox. 1964. "An Analysis of Transformations." *Journal of the Royal Statistical Society: Series B (Methodological)* 26 (2):211–252. URL https://doi.org/10.1111/j.2517-6161.1964.tb00553.x.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Difference-in-Differences with a Continuous Treatment." *Journal of Econometrics* URL https://doi.org/10.3386/w32117. Forthcoming; NBER Working Paper 32117.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/debiased machine learning for treatment and structural parameters." *The Econometrics Journal* 21 (1):C1–C68. URL https://doi.org/10.1111/ectj.12097.

Czajkowski, Mikołaj, Wojciech Zawadzki, Grzegorz Bernatek, and Maciej Sobolewski. 2024. "Assessing the substitutability of mobile and fixed internet: The impact of 5G services on consumer valuation and price elasticity." *Telecommunications Policy* 48 (10):102869. URL https://doi.org/10.1016/j.telpol.2024.102869.

Dunstan, James E. 2023. "The FCC, USF, and USAC: An Alphabet Soup of Due Process Violations." Research Paper 2023/01, Pacific Legal Foundation. URL https://ssrn.com/abstract=4599298.

FCC. 1998. "Third Quarter 1998 Universal Service Contribution Factors Revised and Approved." Federal Register Notice DA-98-1130; CC Docket No. 96-45, FCC. URL https://www.govinfo.gov/content/pkg/FR-1998-07-01/pdf/98-17486.pdf.

———. 2000. "Proposed Second Quarter 2000 Universal Service Contribution Factor." Memorandum DA-00-517A1, FCC. URL https://docs.fcc.gov/public/attachments/DA-00-517A1.pdf.

———. 2014a. "FCC Form 460: Rural Health Care Program Eligibility and Registration Form." Form 460, FCC. URL https://www.usac.org/rural-health-care/healthcare-connect-fund-program/step-1-determine-eligibility-of-your-site/welcome-to-rhc-connect-fcc-form-460/. Issued and released August 5, 2014.

———. 2014b. "FCC Form 461: Rural Health Care Program Request for Services Form." Form 461, FCC. URL https://www.usac.org/rural-health-care/healthcare-connect-fund-program/step-2-prepare-for-competitive-bidding-request-services/welcome-to-rhc-connect-fcc-form-461/. Issued and released August 5, 2014.

———. 2014c. "FCC Form 462: Healthcare Connect Fund Funding Request Form." Form 462, FCC. URL https://www.usac.org/rural-health-care/healthcare-connect-fund-program/step-4-submit-funding-requests/rhc-connect/. Issued and released August 5, 2014; user guide updated May 2023.

GAO. 2012. "Telecommunications: FCC Has Reformed the High-Cost Program, but Oversight and Management Could Be Improved." Tech. Rep. GAO-12-738, GAO. URL https://www.gao.gov/assets/gao-12-738.pdf. Report to Congressional Requesters.

———. 2020. "Telecommunications: FCC Should Enhance Performance Goals and Measures for Its Program to Support Broadband Service in High-Cost Areas." Tech. Rep. GAO-21-24, GAO. URL https://www.gao.gov/assets/gao-21-24.pdf. Report to the Chairman, Committee on Energy and Commerce, House of Representatives.

Gilroy, Angele A. 2011. "Universal Service Fund: Background and Options for Reform." CRS Report RL33979, Congressional Research Service. URL https://crsreports.congress.gov/product/details?prodcode=RL33979.

Hazlett, Thomas W, Ben Schwall, and Scott Wallsten. 2019. "The educational impact of broadband subsidies for schools under E-rate." *Economics of Innovation and New Technology* 28 (5):483–497. URL https://doi.org/10.1080/10438599.2018.1527554.

Niskanen, William A. 1968. "The peculiar economics of bureaucracy." *The American Economic Review* 58 (2):293–305.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics* 37 (2):187–204. URL https://doi.org/10.1080/07350015.2016.1227711.

Rabbani, Maysam. 2024a. "Dollars and megabits: a comparative analysis of Telecom and Healthcare Connect Fund." *Information Economics and Policy* 67:101082. URL https://doi.org/10.1016/j.infoecopol.2024.101082.

———. 2024b. "Internet price, speed, and disparity: The case of rural healthcare providers in the United States." *Telecommunications Policy* 48 (2):102674. URL https://doi.org/10.1016/j.telpol.2023.102674.

U.S. Senate Subcommittee on Communications, Media, and Broadband. 2025. "The State of Universal Service." Hearing before the Subcommittee on Communications, Media, and Broadband of the Committee on Commerce, Science, and Transportation, U.S. Senate, 118th Congress, S. Hrg. 118-535, May 11, 2023. URL https://www.congress.gov/118/chrg/CHRG-118shrg58143/CHRG-118shrg58143.pdf.

USAC. 2024. "Universal Service Fund Annual Report 2024." Tech. rep., USAC, Washington, DC. URL https://www.usac.org/about/reports-orders/annual-report/.

USAC. 2025. "Rural Health Care (RHC) Program: Consortia Submitting the FCC Form 461 and 462." Guide, USAC. URL https://www.usac.org/wp-content/uploads/rural-health-care/documents/Consortia-FCC-Form-461-and-462-Tip-Sheet_Final_01.28.2026.pdf. Revised January 2026.

Wallsten, Scott. 2011. "The Universal Service Fund: What Do High-Cost Subsidies Subsidize?" Tech. rep., Technology Policy Institute. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1927933.

Zolnierek, James. 2008. "Loop density and telephone company cost: Panel data evidence." *Telecommunications Policy* 32 (3-4):262–272. URL https://doi.org/10.1016/j.telpol.2008.01.003.