Partnerships in Urban Mobility: Incentive Mechanisms for Improving Public Transit Adoption

Auyon Siddiq* Christopher S. Tang* Jingwei Zhang*

Problem definition: Due to a prolonged decline in public transit ridership over the last decade, transit agencies across the United States are in financial crisis. To entice commuters to travel by public transit instead of driving personal vehicles, municipal governments must address the "last mile" problem by providing convenient and affordable transportation between a commuter's home and a transit station. This challenge raises an important question: Is there a cost-effective mechanism that can improve public transit adoption by solving the last mile problem?
Academic / Practical Relevance: In this paper, we present and analyze two incentive mechanisms for increasing commuter adoption of public transit. In a *direct* mechanism, the government provides a subsidy to commuters who adopt a "mixed mode", which involves combining public transit with hailing rides to/from a transit station. The government funds the subsidy by imposing congestion fees on personal vehicles entering the city center. In an *indirect* mechanism, instead of levying congestion fees, the government secures funding for the subsidy from the private sector. We examine the implications of both mechanisms on relevant stakeholders. These two mechanisms are especially relevant because several jurisdictions in the U.S. have begun piloting incentive programs in which commuters receive subsidies for ride-hailing trips that begin or end at a transit station.
Methodology: We present a game-theoretic model to capture the strategic interactions among five self-interested stakeholders (commuters, public transit agency, ride-hailing platform, municipal government, and local private enterprises).
Results: By examining equilibrium outcomes, we obtain three key findings. First, we characterize how the optimal interventions associated with the direct or the indirect mechanism depend on: (a) the coverage level of the public transit network; (b) the public transit adoption target; and (c) the relative strength of commuter preferences between driving and taking public transit. Second, we show that the direct mechanism cannot be budget neutral without undermining commuter welfare. However, when the public transit adoption target is not too aggressive, we find that the indirect mechanism can increase both commuter welfare and sales to the private sector partner while remaining budget neutral. Finally, we show that, although the indirect mechanism restricts the scope of government intervention (by eliminating the congestion fee), it can dominate the direct mechanism by leaving all stakeholders better off, especially when the adoption target is modest.
Managerial Implications: Our findings offer cost-effective prescriptions for improving urban mobility and public transit ridership.

Keywords: public transit, public-private partnerships, subsidies, incentives, Mobility as a Service (MaaS).

*UCLA Anderson School of Management, 110 Westwood Plaza, Los Angeles, California, 90095.

1 Introduction

The mission of the United States Federal Transit Administration (FTA) is to "*enhance citizens' mobility, accessibility, and economic well-being through the development and management of public transport services*". However, over the last decade, the FTA's mission has become increasingly threatened by a prolonged decline in public transit ridership across the United States. From 2008 to 2017, per capita ridership on buses, subways, and commuter trains saw a 4% drop in San Francisco, a 5% drop in New York City, and over a 25% drop in both the Los Angeles and Washington DC areas (The Economist, 2019). Moreover, this decline is nationwide: bus ridership has decreased from 5.6 billion trips in 2008 to 4.7 billion trips in 2018 (Kamp, 2020). The widespread decline in public transit ridership is believed to be driven by multiple factors, including lower gasoline prices, changing demographic patterns within cities, and the rise of alternative modes of transportation, such as ride-hailing (Mallett, 2018). As a result of declining fare revenue, over 98% of the 2,200 public transit agencies in the U.S. are in financial crisis (Federal Transit Administration, 2017). Covering these losses is challenging for many municipalities due to limited funding, as well as other competing priorities, such as public safety, education, and affordable housing.

To combat declining ridership and revenue, cities can certainly scale back public transit services. However, doing so would undermine the mission of the FTA for the following reasons. Reducing access to public transit can hinder mobility, especially for those who cannot afford alternative modes of transportation (e.g., private vehicles, ride-hailing services, and taxis). Moreover, the American Public Transit Association (APTA) has found evidence suggesting that investments in public transit can generate economic returns by lifting business sales and promoting job growth (APTA, 2019c). In addition to the economic benefits, public transit plays a vital role in reducing carbon emissions and alleviating traffic congestion (APTA, 2019c).[1] Some cities have offered free bus rides to increase ridership (e.g., in 2019, Lawrence, Massachusetts and Olympia, Washington began offering free bus rides to improve mobility for the poor and the elderly (Kamp, 2020)); however, the long-term financial viability of fully subsidizing bus fares is questionable. Therefore, there is an urgent need for municipalities to develop cost effective – ideally, budget neutral[2] – solutions for increasing transit ridership.

A major barrier to increasing public transit utilization is the "last mile" problem – a challenge caused by the lack of convenient and affordable transit services between an individual's home and a transit station (APTA, 2019b; LA Metro, 2016). Despite increased investments in public transit – total inflation-adjusted funding in the U.S. increased from \$60 to \$72 billion between 2007 and 2017

[1]Single-occupancy vehicles emit 1.5 times the CO_2 emissions of buses, and four times the emissions of subways (Hodges et al., 2010). With respect to traffic, public transit has been found to be critical in reducing congestion during peak hours (Anderson, 2014).

[2]The notion of delivering public services in a budget neutral manner has recently been proposed by the Centers for Medicare and Medicaid Services; the reader is referred to CMS (2018) for details.

(APTA, 2019a) – commuters continue to eschew public transit in favor of personal vehicles, in part due to the last mile problem. Indeed, in 2017, over 85% of US workers used personal vehicles to commute, with less than 6% relying on public transit (Bureau of Transportation Statistics, 2018). Further, while offering subsidized parking near transit stations may appear to be an attractive solution to address the last mile problem, such programs are unlikely to be feasible on a large scale due to limited parking near transit stations, as well as the high cost of building, maintaining, and subsidizing parking spaces. For example, in Los Angeles County, only 24,000 subsidized parking spaces are available, to support approximately 3.6 million daily commuters (Linton, 2016; US Census Bureau, 2018).

To address the last mile problem, various municipal governments are forming partnerships with private transportation companies. For example, in 2019, Los Angeles partnered with the ride-hailing platform Via to launch a pilot project that offers commuters subsidized rides to and from public transit stations. The subsidy is supported by a grant from the Federal Transit Administration (Los Angeles County Metropolitan Transportation Authority, 2019). Several other major cities (e.g., Atlanta, Austin, Detroit, Philadelphia, and Tampa) have also begun testing the use of ride-hailing subsidies as a potential solution to closing the last mile gap (Schwieterman and Livingston, 2018; APTA, 2019e). These subsidy programs have only recently become possible due to advances in information technology, which now enable on-demand ride-hailing, real time tracking of passenger locations, and online/mobile fare payments.[3] More generally, the integration between public transit systems and ride-hailing is accelerating – for example, in early 2020, Uber announced a new in-app feature that coordinates trip drop-off times with train schedules (Uber, 2020). This novel integration between private transportation services and public transit opens the door to new forms of mutually beneficial partnerships.[4]

In this paper, we investigate two incentive mechanisms that aim to increase public transit ridership by addressing the last mile problem. Both mechanisms exhibit the following characteristics:

1. **Both mechanisms have a common feature.** Both mechanisms rely on a strategic partnership between a public transit agency and a private ride-hailing platform. Specifically, in both mechanisms, commuters receive a subsidy for adopting a *mixed mode* of transportation, in which commuters use the ride-hailing service to travel the last mile distance ("x") between their homes and a transit station, and use public transit to travel between the transit station and a final destination (e.g. city center). In this setting, x can be interpreted as a measure of

[3]Information technology has been successfully leveraged to help commuters plan and pay for multi-modal trips within a mobile app; examples include Whim in Helsinki, Citymapper in London, Moovel in Germany, UbiGo in Gothenburg, TAP in Los Angeles, Clipper in San Francisco, and Opal in Sydney (Goodall et al., 2017; Cole, 2017).

[4]There is also evidence to suggest that ride-hailing may be contributing to the decline in public transit ridership, and adding to traffic congestion in city centers (Brown, 2020). Partnerships between public transit agencies and private ride-hailing platforms may facilitate cooperation instead of competition, which can be beneficial to both parties (i.e., increased transit ridership in the city center and increased demand for ride-hailing in suburban areas).

the *coverage*[5] of the public transit system: the last mile x is small (large) if commuters have convenient (inconvenient) access to transit stations from their homes.

2. **Both mechanisms share a common goal.** Both mechanisms are intended to increase the adoption of the mixed mode by a constant factor ("β").

The key differences between the mechanisms are: (a) the role of the local government; and (b) the source of funding for the ride-hailing subsidy. To elaborate, the first incentive program we consider is *direct* in the sense that the the government offers a subsidy ("s") directly to each commuter as an incentive for adopting the mixed mode. To defray the cost associated with the subsidy, the government charges a congestion fee ("e") to commuters who instead choose to drive a personal vehicle into the city.[6] Therefore, under the "direct mechanism" (denoted by [D]), the government uses two levers to entice commuters to adopt the mixed mode: the ride-hailing subsidy s, and the congestion fee e.

Unlike the direct mechanism [D] that relies on congestion fees collected from commuters who drive, the local government secures funding for the subsidy ("z") from the private sector in the second mechanism. This mechanism is *indirect* because the government's role is restricted to facilitating the transfer of the subsidy from a private enterprise to commuters. This scheme is motivated by the fact that many municipalities may not have sufficient funds to further subsidize public transit. To relieve this financial burden, Cole (2017) proposes the following partnership between a municipal government and a private enterprise. The government and public transit agency develop a mobile app to process different transactions, including fares collected from the commuters and subsidies provided by the private enterprise. The private enterprise benefits from higher transit ridership through an expanded customer base. Here, the customer base may represent foot traffic at stores near the transit station (e.g., coffee shops, bakeries, health clinics, spas, hair salons), or number of views on an in-app advertisement. Under this arrangement, the private enterprise can recover the cost associated with the subsidy from the extra revenue derived from the increased demand. In contrast to the direct mechanism, this "indirect mechanism" (denoted by [I]) involves just a single lever: the subsidy z, provided by the private enterprise.

Mechanism [I] is motivated by a growing interest among transit agencies in leveraging private sector partnerships to subsidize public transit trips. For example, since early 2020, commuters in Miami-Dade County can earn loyalty points (i.e., subsidies) by taking public transit, which can

[5] In the context of public transit, "coverage" broadly refers to the accessibility of transit stations by the general population (i.e., commuters in our context) (Tomer et al., 2011).

[6] Congestion fees have been adopted by many municipal governments, including Singapore, Hong Kong, London, Milan, and Stockholm. Singapore was the first to introduce congestion pricing as a tool to control traffic volume, where citizens pay for fees when they enter city center areas (Development Asia, 2018). London introduced congestion fees in 2003, where there is a charge for entering London's congestion charging zone (13 square miles) between 7 a.m. and 6 p.m. on weekdays (Transport for London, 2019). Since the introduction of congestion fees, traffic congestion and private vehicle usage in London has dropped by 25% and 39%, respectively (Badstuber, 2018). Similarly, in Milan, public ridership increased 12.5% from 2007 to 2013 (Croci, 2016).

be redeemed for discounts on future trips through the transit agency's mobile app (known as "Go Miami-Dade Transit") (Corselli, 2020). These loyalty points are funded by local businesses who post in-app video ads within the mobile app. The National Hockey League (NHL) and the Seattle Monorail formed a partnership in which NHL will subsidize rides on the Monorail for fans attending hockey games (Condor, 2020). Both of partnerships can be viewed as a private enterprise (i.e., a single aggregate entity) subsidizing public transit trips in exchange for an expanded audience, which capture the main characteristics of mechanism [I].

1.1 Outline and contributions

Our intent is to investigate the impact of both subsidy mechanisms on multiple stakeholders: a municipal government, suburban commuters, city dwellers, a public transit agency, a ride-hailing platform, and a private enterprise. To do so, we examine the following questions:

1. How does each mechanism perform in terms of: (a) operating cost (borne by the government); (b) commuter welfare; (c) the public transit agency's revenue; (d) the ride-hailing platform's revenue; and (e) the private enterprise's profit?

2. How does the last mile distance (or coverage) x and the mixed mode adoption target β affect: (a) the optimal direct subsidy s^* and congestion fee e^* in mechanism [D], and (b) the optimal indirect subsidy z^* in mechanism [I] ?

3. Under what conditions, if any, should the municipal government adopt mechanism [I] over mechanism [D]?

We examine the above questions by developing a game-theoretic model that captures the interactions among all stakeholders. Specifically, in mechanism [D], we identify the optimal subsidy level s^* and congestion fee e^* that minimizes the government's net spending (subsidy cost less the revenue generated by the congestion fee), subject to improving adoption of the mixed mode by a factor of at least β. In addition, we restrict the set of feasible subsidies and congestion fees to those that satisfy participation constraints associated with commuters, the ride-hailing platform, the public transit agency, and the private enterprise. We also determine the optimal subsidy z^* under mechanism [I], subject to the same set of stakeholder participation constraints.

Our key findings are summarized as follows:

1. **The optimal incentive under mechanism [D] depends on the last mile distance x.** Specifically, the optimal incentive entails a large congestion fee and a small ride-hailing subsidy when the transit coverage is high (i.e. when the last mile distance x is small). Conversely, when transit coverage is low (i.e., x is large), this prescription is reversed: the optimal incentive relies on aggressive ride-hailing subsidies and relatively smaller congestion

fees. Moreover, we find that whether the optimal subsidy s^* increases or decreases in the last mile distance x depends on the commuter's relative preference between driving and taking public transit.

2. **Budget neutrality is not attainable under mechanism [D]:** It is not possible for the government to fully recoup the cost of the subsidy through congestion fees without decreasing total commuter welfare.

3. **Mechanism [I] can *dominate* mechanism [D] under certain conditions.** Despite the government having access to two levers in mechanism [D] (the subsidy s and the congestion fee e) and only a single lever in mechanism [I] (the subsidy z), mechanism [I] can *dominate* mechanism [D], but only when the adoption target β is modest. Specifically, when β is modest, the commuters, city dwellers, public transit agency, ride-hailing platform, and private enterprise are all better off under mechanism [I].

The remainder of the paper is organized as follows. In §1.2, we discuss related literature. In §2, we present a general model for analyzing both mechanisms. In §3, we present mechanism [D] and analyze the optimal incentive scheme (e^*, s^*). In §4, we present mechanism [I], analyze the optimal subsidy z^*, and compare the performance of both mechanisms. In §5, we consider variants of mechanisms [D] and [I] that address the trade-off between operating cost and commuter welfare. In §6, we conclude and discuss potential future research directions.

1.2 Related literature

Our paper is related to three streams of literature: government subsidy programs, public-private partnerships, and budget neutral mechanisms.

Government subsidy programs. Within the operations literature, there is a growing focus on subsidy programs that promote the production or adoption of socially beneficial goods or services. This work can be categorized into subsidies for producers (e.g., farmers, manufacturers, healthcare providers) and consumers. Subsidies for producers have been studied in the context of healthcare (Taylor and Xiao, 2014; Levi et al., 2017; Aswani et al., 2019), green technology (Ma et al., 2019; Bansal and Gangopadhyay, 2003; Alizamir et al., 2016), and agriculture (Alizamir et al., 2019; Akkaya et al., 2019). Our paper is closer to the literature on consumer-facing subsidies. In the home appliances industry, Yu et al. (2018) determine whether the government should subsidize consumers only, manufacturers only, or both, where the goals of the government are to improve manufacturer profit and consumer welfare in rural areas. Xiao et al. (2019) examine the impact of the same home appliance subsidy program, which they conclude improves both affordability and accessibility for rural customers. In the context of solar energy technology, Chemama et al. (2019) compare the effectiveness of static and dynamic consumer subsidies on supplier behavior.

Our paper differs from this existing literature in that we investigate whether the subsidy program can be budget neutral from the perspective of the subsidy provider.

Previous work in the operations literature on public transit subsidies is sparse. Lodi et al. (2016) consider a setting where operation of the public transit system is outsourced to the private sector, and the government provides a subsidy to offset the operating cost. Yang and Lim (2018) use a field experiment to show that temporarily subsidizing public transit can lead to long-term changes in commuter behavior. The paper that is most similar to ours in this line of research is by Xiao and Zhang (2014), who also consider congestion fees and transit subsidies simultaneously. They focus on the impact that commuter heterogeneity in value-of-time has on the optimal design of congestion fees, and show that transit subsidies can offset the loss in commuter welfare due to the congestion fee. Our work is different in that we focus on the role that congestion fees and transit subsidies can play in improving public transit ridership.

Public-private partnerships. Our work contributes to the modeling of public-private partnerships (PPP), which refers to a private sector partner "financing, constructing, and managing a project in return for a promised stream of payments directly from government or indirectly from users over the projected life of the project or some other specified period" (Weimer and Vining, 2017). Our paper belongs to the stream of PPP literature where the government has all of the bargaining power. In the existing literature, this is often modeled as either a principal-agent problem or a Stackelberg game, where the government is the principal/leader and the firm is the agent/follower; applications include healthcare (So and Tang, 2000; Lee and Zenios, 2012; Gupta and Mehrotra, 2015; Guo et al., 2019; Aswani et al., 2019), disaster management (Guan and Zhuang, 2015; Guan et al., 2018) and risk management (Bakshi and Gans, 2010). With respect to transportation, existing work on public-private partnerships has primarily focused on infrastructure. Lodi et al. (2016) address the government's incentive design problem when management of the public transit service is outsourced to private operators. Gagnepain and Ivaldi (2002) examine the effect on social welfare under different types of contracts between the regulator and the public transit operator. Hansson (2010) considers a multi-principal setting, where the local, regional and county governments interact to regulate public transit procurement. In contrast to these papers, the private partners play a different role in our work, namely, they provide a complementary transportation service (in the case of the ride-hailing platform) and funding for subsidies (in the case of the partner enterprise).

Previous work has also considered settings where the public-private partnership is formed through negotiation. In the transportation context, Kang et al. (2013) investigate royalty bargaining associated with underground railway station construction, and Wang and Zhang (2016) examine road pricing of transportation networks with both public and private roads. Other application areas include natural resource development (Anandalingam, 1987), public procurement (Gur et al., 2017; Saban and Weintraub, 2019) and global supply chain management (Cohen et al., 2018; Cho et al., 2019; Cohen and Lee, 2020).

Budget neutral mechanisms. This paper also contributes to the literature on budget neutral policies. This work has primarily appeared in the public policy literature, and has addressed issues such as social security reform (Burkhauser and Smeeding, 1994), environmental taxation (Goulder, 1995; Murray and Rivers, 2015), and fiscal policy (Correia et al., 2013; D'Acunto et al., 2016). Within the operations management literature, previous work on budget neutral policies is scarce. Guo et al. (2014) optimize a two-tier queuing system with both a free server and fee-based server, in the setting where the system is self-financed by the costly server. Arifoglu and Tang (2019) develop a budget neutral incentive mechanism for coordinating a decentralized influenza vaccine supply chain.

With respect to transportation settings, most existing papers that focus on budget neutrality are based on schemes under which subsidies/rebates are funded by congestion fees, similar to mechanism [D] in our work (see, e.g., Guo and Yang (2010); Nie and Liu (2010); Chen and Yang (2012); Xiao and Zhang (2014)). Our paper differs in that we also consider an indirect mechanism [I], where the government attains budget neutrality by obtaining funding from a private enterprise, instead of imposing congestion fees. Further, we compare mechanisms [D] and [I] in terms of their impact on the relevant stakeholders.

2 Model Preliminaries

A unit mass of commuters are located a (last mile) distance $x > 0$ from a transit station. All commuters must travel to a city center that is located an additional distance of 1 beyond the transit station. For tractability, we consider a parsimonious model where commuters choose between two modes of travel:[7] *driving* or a *mixed mode*. As depicted in Figure 1, x represents the length of the "last mile" that is not covered by public transit. Hence, commuters who choose the driving mode will drive a personal vehicle for a distance of $1 + x$ to the city center. However, commuters who choose the mixed mode will first travel a distance of x to the transit station via a ride-hailing service, and then travel the remaining distance of 1 by public transit.

2.1 Base Model

We first describe our base model in the absence of any subsidies or congestion fees. In later sections, we extend our base model to incorporate the two incentive schemes. We note here that our focus

[7] We focus on suburban commuters for whom walking to the transit station is prohibitively costly. Accordingly, we exclude the mixed mode that combines walking and public transit. Similarly, we exclude the case where commuters combine electric scooters or bicycles and public transit because (a) commuters' comfortable biking distance is limited (Rastogi and Krishna Rao, 2003) and (b) less than 4% of commuters use scooters or bicycles to get to/from transit stations (Federal Highway Administration, 2017)). We also exclude the mixed mode that combines driving and public transit, due to limited parking spaces near transit stations in the U.S. Commuters who exclusively use ride-hailing to commute are also not considered here(only 0.2% of suburban residents commute via taxi or ride-hailing services (Federal Highway Administration, 2017)).

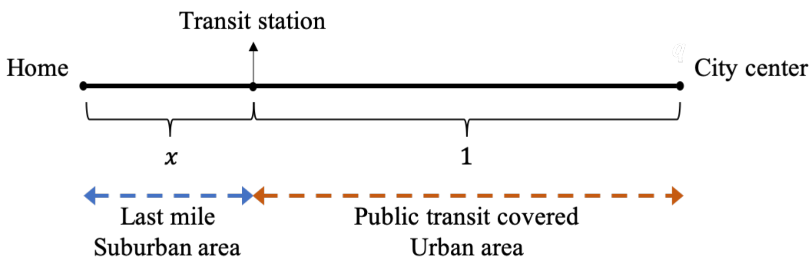

Figure 1: Travel distance of a commuter.

throughout the paper is on daily commuters who live in suburban areas outside the coverage area of the public transit system; therefore, all aspects of the model (e.g., transit ridership, operating cost, ride-hailing platform revenue) are defined with respect to this group of commuters only.

Commuter utility. Each commuter obtains a "base" utility M for commuting to the city center. A commuter has valuation (or willingness-to-pay) V per unit distance traveled via ride-hailing or public transit, or δV if she drives.[8] We assume $\delta > 1$ to reflect a higher intrinsic utility for driving.[9] To capture heterogeneity among commuters, we assume $V \sim U[0,1]$.[10] (In Appendix C, we consider an alternative setting in which V is deterministic and identical for all commuters. Instead, we consider commuter heterogeneity in terms of x by assuming that x (i.e., the distance between a commuter's home and the transit station) is uniformly distributed. For this alternative setting, we obtain the same structural result.)

A commuter will travel a distance of $1 + x$ if she chooses the driving mode of transportation. Similarly, under the mixed mode, the commuter will first travel a distance of x via ride-hailing and the remaining distance of 1 via public transit. In our model, we assume a commuter's cost (or fare) for driving, ride-hailing, and taking public transit are denoted by d, r, and p, respectively. We assume throughout that $r > d > p > 0$, which is supported empirically.[11] Therefore, the utilities associated with driving and the mixed mode in the absence of any incentives, denoted by U_d^0 and

[8]Empirical evidence suggests that a commuter's willingness-to-pay for transportation is increasing in the travel distance (Van Ommeren et al., 2000; Jou et al., 2012). For tractability, we assume a linear relationship between willingness-to-pay and distance, given by the parameter V.

[9]Recent survey data has shown that commuters generally prefer driving to public transit and ride-hailing services (Zhu and Fan, 2018). Our main results also persist in a model where commuters prefer ride-hailing to public transit, so that the additional value associated with ride-hailing is $l \cdot V$, where $\delta > \ell \geq 1$. For ease of exposition, we assume $\ell = 1$ throughout.

[10]Commuter valuation for each travel mode can vary based on demographic features such as age, income level, or health conditions (Zhu and Fan, 2018). For simplicity, we assume the valuation $V \sim U[0,1]$.

[11]The cost of driving is estimated to be \$0.76/mile, assuming a mileage of 10,000 miles per year (American Automobile Association, 2018). Ride-hailing is estimated to cost more than \$1.07/mile (Fare Estimater, 2019). The cost of commuting by public transit is typically much lower; for example, the cost of public transit in Los Angeles is estimated to be \$0.2/mile, assuming an average commute distance of 16 miles (Leonard, 2019).

U_m^0, respectively, are given by

$$U_d^0 = M + (x+1)\delta V - (x+1)d, \tag{1a}$$

$$U_m^0 = M + (x+1)V - (rx+p). \tag{1b}$$

Travel mode demand. It follows from (1) that a commuter will adopt the mixed mode if and only if $U_m^0 \geq U_d^0$. We assume the base value M is large enough such that $U_d^0 \geq 0$ and $U_m^0 \geq 0$. Therefore, $U_m^0 \geq U_d^0$ if and only if:

$$V \leq v^0 \equiv \frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}. \tag{2}$$

Let D_d^0 and D_m^0 denote the demand for driving and the mixed mode, respectively. Because $V \sim U[0,1]$, we can apply (2) to show that:

$$D_d^0 = \int_{v^0}^{1} 1 \cdot dV = 1 - \frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}, \tag{3a}$$

$$D_m^0 = \int_{0}^{v^0} 1 \cdot dV = \frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}. \tag{3b}$$

Note that D_m^0 represents the additional public transit ridership generated by the mixed mode commuters. Because $r > d > p > 0$, it is straightforward to verify that D_m^0 is decreasing in x, which implies that commuters that live far from the transit station are more likely to drive and less likely to adopt the mixed mode. Further, D_m^0 also measures the "reduction" in traffic congestion, because $D_m^0 = 1 - D_d^0$.

Let $[a]^+ = \max\{a, 0\}$. To exclude the degenerate cases where *all* commuters adopt the same travel mode, we assume that $x \in (\underline{x}, \bar{x})$, so that $D_m^0 > 0$ and $D_d^0 > 0$, where $\underline{x} = \left[\frac{(d-p)-(\delta-1)}{(\delta-1)+(r-d)}\right]^+$ and $\bar{x} = \frac{d-p}{r-d}$. We also assume that the public transit and the ride-hailing platform have sufficient capacity to accommodate demand for the mixed mode.[12]

Commuter welfare. Because $V \sim U[0,1]$, the commuter welfare in the base model is given by

$$W^0 = \int_{0}^{v^0} U_m^0 dV + \int_{v^0}^{1} U_d^0 dV. \tag{4}$$

Transit agency revenue. By noting that the unit fare of public transit is p, the revenue generated by the public transit agency from these commuters under the mixed mode demand D_m^0 is:

$$\Pi_p^0 = p \cdot D_m^0. \tag{5}$$

Ride-hailing platform revenue. Recall that each mixed mode commuter will travel a distance of x via ride-hailing to the transit station, at a unit fare of r. The revenue generated by the

[12] The average vehicle occupancy rate of public transit in 2017 was less than 30% (APTA, 2019d). Ride-hailing services also have ample capacity: the average idle rate of Uber drivers is 30-40% (Currie, 2018; Brown, 2020).

ride-hailing platform from these commuters under the mixed mode demand D_m^0 is:

$$\Pi_r^0 = r \cdot x \cdot D_m^0. \tag{6}$$

Private enterprise profit. Consider a private enterprise who sells to public transit commuters. This private enterprise may be a business that sells goods or services to passengers that pass through a transit station, or one that places advertisements within a transit system's mobile app or physical walkways. We model the enterprise profit as $(k-c)[1-k\alpha]^+ D_m^0$, where k and c represent the unit price and cost, respectively, and $[1-k\alpha]^+$ represents the proportion of mixed mode commuters that purchase the product. Note that $\alpha > 0$ is the commuters' price sensitivity. To avoid the trivial case where the enterprise's optimal profit is non-positive, we assume $\alpha c < 1$.[13] Then, for any given mixed mode demand D_m^0, the optimal retail price in the base model is $k_b = \frac{\alpha c+1}{2\alpha}$ so that the corresponding optimal profit is equal to $(k_b - c)[1-k_b\alpha]^+ D_m^0 = \frac{(1-\alpha c)^2}{4\alpha} D_m^0$. For ease of exposition, we let $K \equiv \frac{(1-\alpha c)^2}{4\alpha}$ to denote the profit margin so that the corresponding total profit can be expressed as:

$$\Pi_s^0 = K \cdot D_m^0. \tag{7}$$

2.2 A Generalized Mechanism: Government Subsidy, Congestion Fees and Private Subsidy

By using the quantities as defined above, we now present our model that incorporate the aforementioned incentives. To avoid repetition, we shall present a generalized mechanism that combines three interventions: (1) a congestion fee e charged to each commuter who travels by a personal vehicle; (2) a subsidy s for each commuter who takes the mixed mode, paid for by the government; and (3) a subsidy z for each commuter who takes the mixed mode, paid for by the private enterprise. We denote this generalized mechanism by (e, s, z). When $z = 0$, this mechanism reduces to mechanism [D]. Also, when $e = s = 0$, the mechanism simplifies to mechanism [I]. To isolate the effect of (e, s) under mechanisms [D] and the effect of s under mechanism [I], we shall analyze mechanism [D] in §3 and mechanism [I] in §4 as two separate mechanisms. (Based on our knowledge, most transit agencies adopt either mechanism [D] or [I], but no both. However, we shall analyze the hybrid mechanism that combines both mechanisms [D] and [I] in Appendix A (i.e., when (e, s) and z are permitted to be non-zero) and show that the structural results continue to hold.)

The sequence of events is as follows. First, the government chooses whether to adopt mechanism [D] or mechanism [I]. Second, the government sets the incentives that correspond to the chosen mechanism. Specifically, the government sets the congestion fee e and government subsidy s if mechanism [D] is chosen, or the private subsidy z if mechanism [I] is chosen. Next, once (e, s) or z is set, the private enterprise selects the unit price k to maximize its profit. Finally, each commuter

[13]Note that if $\alpha c \geq 1$, then $(k-c)[1-k\alpha]^+ D_m^0 \leq 0$ for any $D_m^0 \geq 0$ and $k \geq 0$.

chooses to commute by either driving or adopting the mixed mode.

By considering (1) along with the incentive (e, s, z), commuter utilities associated with different modes of transportation can be written as:

$$U_d(e, s, z) = U_d^0 - e, \tag{8a}$$

$$U_m(e, s, z) = U_m^0 + s + z, \tag{8b}$$

where U_d^0 and U_m^0 are given in (1).[14]

Hence, a commuter will adopt the mixed mode under incentive (e, s, z) if and only if $U_m(e, s, z) \geq U_d(e, s, z)$, or equivalently, if her valuation V satisfies:

$$V \leq v(e, s, z) \equiv \min\left\{1, \frac{(d-p) - (r-d)x + e + s + z}{(\delta - 1)(x+1)}\right\}. \tag{9}$$

Akin to the base case without incentives, the demand for each transportation mode under the incentive (e, s, z) is given by

$$D_d(e, s, z) = \left[1 - \frac{(d-p) - (r-d)x + e + s + z}{(\delta - 1)(x+1)}\right]^+, \tag{10a}$$

$$D_m(e, s, z) = 1 - \left[1 - \frac{(d-p) - (r-d)x + e + s + z}{(\delta - 1)(x+1)}\right]^+. \tag{10b}$$

Note that if $(e, s, z) = (0, 0, 0)$, then $v(e, s, z)$, $D_d(e, s, z)$, and $D_m(e, s, z)$ reduce to their base values v^0, D_d^0, and D_m^0 as stated in §2.1, respectively.

Performance metrics. Next, we define the metrics by which we evaluate the performance associated with the incentive (e, s, z). Because the government offers a subsidy s to each of the $D_m(e, s, z)$ mixed mode commuters and collects a fee e from each of the $D_d(e, s, z)$ commuters who drive, and meanwhile the government passes on a subsidy z from the partner enterprise to each mixed mode commuter, the total cost for the government to operationalize the incentive (e, s, z) is:

$$\begin{aligned} C(e, s, z) &= (s + z)D_m(e, s, z) - eD_d(e, s, z) - zD_m(e, s, z) \\ &= sD_m(e, s, z) - eD_d(e, s, z). \end{aligned} \tag{11}$$

Note that when $e = s = 0$, there is no operating cost $C(0, 0, z) = 0$. Similar to the base case, the total welfare accrued to all commuters is:

$$W(e, s, z) = \int_0^{v(e,s,z)} U_m(e, s, z)dV + \int_{v(e,s,z)}^1 U_d(e, s, z)dV, \tag{12}$$

the public transit agency's revenue generated from these commuters is:

$$\Pi_p(e, s, z) = p \cdot D_m(e, s, z), \tag{13}$$

[14]Observe from (1) that the utilities U_d^0 and U_m^0 are based on parameters x, d, r and p. Hence the lump-sum congestion fee e and subsidy s and z in (8) can also be expressed as a percentage fee and subsidy through an appropriate scaling of those parameters.

and the ride-hailing platform's revenue generated from these commuters is:

$$\Pi_r(e,s,z) = r \cdot x \cdot D_m(e,s,z). \tag{14}$$

Lastly, the private enterprise's profit generated from these commuters is $((k-c)[1-k\alpha]^+ - z)D_m(e,s,z)$. Observe that the optimal price $k_g = \frac{\alpha c+1}{2\alpha}$ so that the corresponding optimal profit is $(K-z)D_m(e,s,z)$, where K is the enterprise's profit margin in the base model as defined in §2.1. We assume the profit margin of a participating enterprise K satisfies $K > \underline{K}$ throughout the paper, where $\underline{K} = (d-p)-(r-d)x$, because enterprises with low profit margin would not opt in the subsidy program. Hence, the enterprise's total profit under incentives (e,s,z) is

$$\Pi_s(e,s,z) = (K-z) \cdot D_m(e,s,z). \tag{15}$$

Note that when $(e,s,z)=(0,0,0)$, the metrics $W(e,s,z)$, $\Pi_p(e,s,z)$, $\Pi_r(e,s,z)$, and $\Pi_s(e,s,z)$ reduce to the base values W^0, Π_p^0, Π_r^0, and Π_s^0 (defined in §2.1), respectively.

Next, we present the evaluation criteria for the incentive program (e,s,z). In particular, we assume that (e,s,z) is set according to the following criteria:

I **Minimum operating cost.** Because most municipalities in the U.S. and around the world are budget constrained with respect to public transit,[15] we assume that the government has a strong desire to minimize the total operating cost of the incentive program, but not profit from it, where the government's operating cost is given by $[C(e,s,z)]^+$.

II **Increase transit ridership.** The intent of the incentive (e,s,z) is to increase public transit ridership among commuters by a factor of $\beta > 0$. We assume throughout that β is an exogenous adoption target, and that the incentive (e,s,z) must satisfy:

$$D_m(e,s,z) - D_m^0 \geq \beta D_m^0,$$

where D_m^0 is the baseline mixed mode adoption without any intervention as stated in §2.1. Note that the increase β is equivalent to a decrease in demand for driving; therefore, β may be equivalently interpreted as the target reduction in traffic congestion. For this reason, the constraint above can also be interpreted as improving the well-being of city dwellers. Analogous to the assumption that $x \in (\underline{x}, \bar{x})$ in the base model, we assume throughout that the target β is restricted to the non-degenerate case where $D_d(e,s,z) > 0$ and $D_m(e,s,z) > 0$. To enforce the preceding inequalities, it suffices to assume that $\beta < \bar{\tau}(x) \equiv \frac{(\delta-1)(x+1)}{(d-p)-(r-d)x} - 1$. Note that if $\beta \geq \bar{\tau}(x)$, then upon implementation of incentive (e,s,z), no commuters will drive $(D_d(e,s,z)=0)$, which is unlikely to occur in practice.

[15]Due to an anticipated budget deficit through 2026, the U.S. Department of Transportation is likely to further reduce transfers to states and transit agencies (Kirk and Mallett, 2020). Similarly, Transport for London faces a grant cut from the government of £700m a year (Edwards, 2018).

III **Commuter participation.** To generate public support for the incentive program, commuters should not be worse off collectively: $W(e,s,z) \geq W^0$.

IV **Ride-hailing platform participation.** To form a viable partnership with the ride-hailing platform, the platform's revenue generated from daily commuters should not decrease: $\Pi_r(e,s,z) \geq \Pi_r^0$.

V **Public-transit participation.** To ensure that the public transit agency is not negatively impacted, its revenue generated from the mixed mode should not decrease: $\Pi_p(e,s,z) \geq \Pi_p^0$.[16]

VI **Enterprise participation.** Lastly, to prevent adverse effects on the local economy, it is desirable to ensure that the incentive (e,s,z) does not reduce the private enterprise's profit: $\Pi_s(e,s,z) \geq \Pi_s^0$. Further, in mechanism [I] in particular, to entice participation of the private enterprise, the private subsidy z is set to maximize the private enterprise's profit: $z = \arg\max_z \Pi_s(e,s,z)$.

Observe from criterion II that a feasible incentive program (e,s,z) must satisfy $D_m(e,s,z) \geq D_m^0$. Combine this observation with (13) and (14), criteria IV and V are satisfied (and hence, redundant) immediately if $D_m(e,s,z) \geq D_m^0$. This result is intuitive: when the government increases the mixed mode adoption, both the ride-hailing company and the transit agency will be better off. Therefore, we shall omit criteria IV and V from the incentive design problems in the remainder of the paper for conciseness.

3 Direct Mechanism: Commuter Subsidies and Congestion Fees

In the direct mechanism [D], the local government directly offers a subsidy s to each mixed mode commuter, and charges a congestion fee e to each commuter who drives. Hence, by letting $z = 0$ in the general formulation presented in §2.2, we can determine the commuter utility and the market demand associated with the driving mode and the mixed mode, along with all the performances metrics under mechanism [D]. Figure 2 depicts the relationship among the various stakeholders in this case.

Based on the evaluation criteria I $\sim$ VI (with IV and V being omitted for conciseness), the

[16] Note that although transit agencies receive funding from local governments, they operate under separate budgets (Federal Transit Administration, 2020). Therefore, we model the transit agency as a distinct stakeholder from the local government.

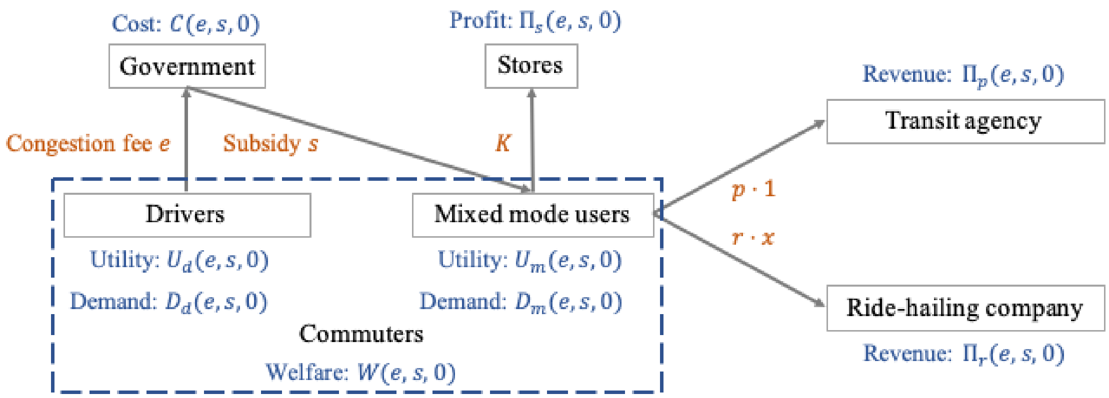

Figure 2: Strategic interactions among stakeholders under the indirect mechanism [D].

optimal incentives e^* and s^* are the solution to the following problem:

$$\text{Mechanism [D]:} \quad \begin{aligned} \min_{e,s\geq 0} \ & [C(e,s,0)]^+ && \text{(16a)} \\ \text{s.t. } & D_m(e,s,0) - D_m^0 \geq \beta D_m^0 && \text{(16b)} \\ & W(e,s,0) \geq W^0 && \text{(16c)} \\ & \Pi_s(e,s,0) \geq \Pi_s^0. && \text{(16d)} \end{aligned}$$

3.1 Direct Mechanism: Optimal Incentive (e^*, s^*)

By solving problem (16), we obtain the optimal incentive (e^*, s^*) under mechanism [D].

Proposition 1. *The optimal congestion fee e^* and the optimal subsidy s^* under mechanism [D] are given by:*

$$\begin{aligned} e^* &= \frac{\beta(\beta+2)((d-p)-(r-d)x)^2}{2(\delta-1)(x+1)}, \\ s^* &= \frac{\beta((d-p)-(r-d)x)(2(\delta-1)(x+1)-(\beta+2)((d-p)-(r-d)x))}{2(\delta-1)(x+1)}. \end{aligned} \tag{17}$$

Further,

(i) The optimal congestion fee e^* strictly increases in β and strictly decreases in x.

(ii) The optimal subsidy s^* strictly increases in β. If δ is large, s^* strictly decreases in x. However, if δ is small, there exists $\tilde{x}$ such that s^* increases on $x < \tilde{x}$ and decreases on $x \geq \tilde{x}$.

Statements (i) and (ii) of Proposition 1 imply that, as the mixed mode adoption target β becomes more aggressive, the commuter subsidy s^* and congestion fee e^* also increase to meet the adoption target, as expected. Additionally, as the last mile distance x increases, the demand for driving also increases, and hence the congestion fee e^* decreases to maintain commuter welfare.

In contrast to the congestion fee e^*, the behavior of the optimal subsidy s^* is not monotonic in the last mile distance x. To see why, let us consider the impact on mechanism [D] when x

increases. As x increases, the mixed mode adoption (before any government intervention) D_m^0 (as stated in (3)) decreases, which generates two competing effects. First, note that the transit ridership constraint can be rewritten as $D_m(e,s,0) \geq (\beta+1)D_m^0$. Therefore, as D_m^0 decreases, the requirement for the mixed mode commuters $D_m(e,s,0)$ becomes less stringent. Consequently, the government can afford to reduce s, which we refer to as the "*adoption effect*." Second, as D_m^0 decreases, commuter welfare decreases, due to fewer commuters receiving the subsidy. To ensure commuters are not worse off, the government needs to increase s, which we refer to as the "*welfare effect*." Hence, whether the optimal subsidy s^* increases or decreases in x depends on which of these two effects dominate.

To examine when one effect dominates the other, first consider the case when commuters strongly prefer driving over transit and ride-hailing (i.e., when δ is large). In this case, the welfare effect is weak (because the mixed mode adoption D_m^0 is already low before any intervention), and the adoption effect dominates – leading s^* to decrease in x. However, when commuters are relatively indifferent between the two modes (i.e., when δ is small), the mixed mode adoption D_m^0 is highly sensitive to small changes in x. As a result, as x increases, demand for the mixed mode drops sharply, which makes the welfare effect dominates – leading s^* to increase in x.[17]

Corollary 1. *There exists a threshold $\hat{x}$ such that $e^* \geq s^*$ in mechanism [D] if and only if $x \leq \hat{x}$. Further, there exists a threshold $\bar{\bar{\delta}} > 1$ such that $s^* > e^*$ in mechanism [D] for all $x \in (\underline{x}, \bar{x})$ if and only if $\delta \geq \bar{\bar{\delta}}$.*

Corollary 1 states that for any adoption target β, the optimal congestion fee e^* is comparatively larger than the optimal subsidy s^* when the last mile distance x is small. To see why, observe that when x is small, the mixed mode is already attractive to commuters, which makes providing a subsidy (per commuter) relatively costly; in this setting, the congestion fee e is the preferred lever for promoting the adoption of the mixed mode, due to its cost efficiency. Conversely, when the last mile distance x is large, more aggressive subsidies are required to meet the mixed mode adoption target. Corollary 1 also implies that if δ is large, then $\hat{x} < \underline{x}$, which yields $s^* > e^*$ for all $x \in (\underline{x}, \bar{x})$. This occurs because when δ is large, most commuters prefer to drive, and so it is sub-optimal for the government to set a high congestion fee, due to its degradation of commuter welfare.

Next, we examine whether the government can attain budget neutrality (or positive revenue) by using congestion fees to offset commuter subsidies under mechanism [D].

Corollary 2. *It is impossible for mechanism [D] to be budget neutral, i.e., $C(e^*, s^*, 0) > 0$. Further, the minimal operating cost $C(e^*, s^*, 0)$ is higher when β is large or when x is small.*

The intuition behind Corollary 2 is as follows. For any budget neutral incentive (e, s) (i.e., where the congestion fee e is selected to fully cover the cost of the subsidy s), it can be shown that the utility

[17]Note that even when δ is small, the welfare effect ceases to dominate the adoption target effect when the last mile distance is large ($x \geq \tilde{x}$), due to low demand for the mixed mode.

loss to drivers exceeds the utility gain to transit riders. In other words, for any subsidy s, there is no congestion fee e that can achieve budget neutrality without a net decrease in total commuter welfare. As a result, the requirement that commuter welfare be non-decreasing in mechanism [D] can only be satisfied through a positive operating cost. Next, to see why $C(e^*, s^*, 0)$ is higher when β is large or when x is small, observe that, in this case, the mixed mode adoption target βD_m^0 on the right hand side of constraint (16b) is large, which requires the mixed mode adoption $D_m(e, s, 0)$ to also be large. Therefore, as more mixed mode commuters receive the subsidy, mechanism [D] incurs a greater cost to the government. We can now summarize our main finding from mechanism [D] as follows.

Remark 1. *The direct mechanism [D] can enable the government to meet the public transit adoption target β. However, this mechanism is costly: attaining budget neutrality is impossible without undermining commuter welfare.*

Note that our commuter utility function implies that commuters are equally sensitive to the subsidy and congestion fee; that is, $|\partial U_m(e, s, 0)/\partial s| = |\partial U_d(e, s, 0)/\partial e|$. Note that this symmetry may not hold if commuters are loss averse. However, we find that the structure of our main results continues to hold under loss aversion as well. Intuitively, if commuters perceive losses from congestion fees more strongly than gains from subsidies, then the government would need to *increase* the subsidy and *decrease* the congestion fee to satisfy the non-decreasing commuter welfare constraint, resulting in a higher total operating cost.[18]

Remark 1 naturally raises the following question: Does there exist an alternate mechanism that can enable the government to attain budget neutrality, without negatively impacting commuter welfare? In the next section, we address this question by examining an alternative incentive program under which the subsidy is funded by a private enterprise, instead of congestion fees.

4 Indirect Mechanism: Funding Subsidies Through a Private Enterprise

Because most municipalities are financially constrained, self-funded or budget neutral projects are preferred. In this section, we present the *indirect* mechanism [I], where the local government secures funding for the subsidy entirely from a private enterprise,[19] and does not charge congestion fees.

[18] Similar to Homonoff (2018), we can extend our model to capture one aspect of loss aversion by defining the commuter utilities associated with the mixed mode and driving as $U_m(e, s, z) = U_m^0 + s + z$ and $U_d(e, s, z) = U_d^0 - \rho \cdot e$, respectively, where $\rho \geq 1$ captures the strength of the commuter's aversion to the congestion fee. The evaluation criteria I – VI can be re-defined accordingly. Using a parallel argument to the proof of Corollary 2, it is straightforward to verify that the optimal operating cost is strictly positive for all $\rho \geq 1$.

[19] As explained in §1, the transit agency can generate revenue from private companies who post in-app video ads via mobile apps, and the indirect mechanism can be operationalized through a mobile app that allows the government to collect fares from riders and subsidies from the private sector.

Note that the indirect mechanism is budget neutral by design. We will therefore focus on the performance of mechanism [I] on the remaining evaluation criteria.

By using the same approach as in §3, we now present our model associated with mechanism [I], which corresponds to the special case of the general model framework presented in §2.2 when we set $(e, s) = (0, 0)$. In other words, by replacing (e, s, z) in the general formulation with $(0, 0, z)$, we can determine the performance metrics under mechanism [I]. The interactions among all stakeholders are depicted in Figure 3.

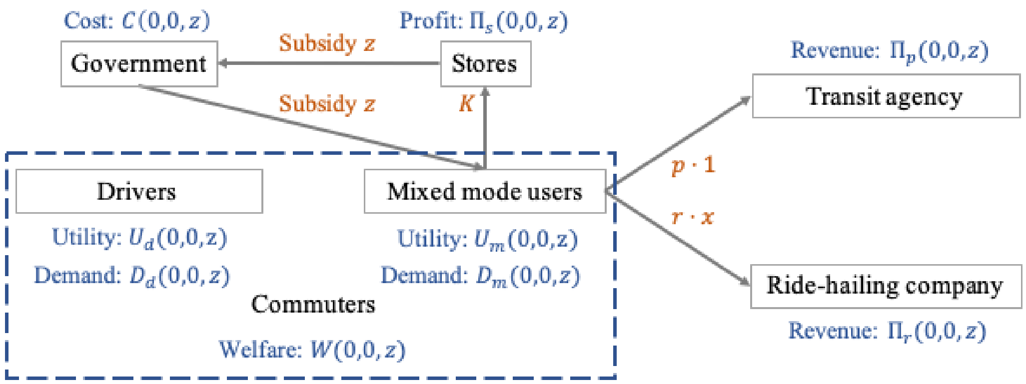

Figure 3: Strategic interactions among stakeholders under the indirect mechanism [I].

Because the viability of mechanism [I] depends on participation from the private enterprise, we assume the private subsidy z is set to maximize the enterprise's total profit. Based on the evaluation criteria I – VI, the government's incentive design problem associated with subsidy z under mechanism [I] can be formulated as the following optimization problem:

$$\text{Mechanism [I]:} \quad \begin{aligned} &\min_{z \geq 0} \; [C(0,0,z)]^+ && \text{(18a)} \\ &\text{s.t.} \; D_m(0,0,z) - D_m^0 \geq \beta D_m^0 && \text{(18b)} \\ &\quad W(0,0,z) \geq W^0 && \text{(18c)} \\ &\quad \Pi_s(0,0,z) \geq \Pi_s^0 && \text{(18d)} \\ &\quad z = \arg\max_z \Pi_s(0,0,z) && \text{(18e)} \end{aligned}$$

Note that $C(0, 0, z) = 0$ for all z, by the definition of $C(e, s, z)$ as given in (11). Therefore, under mechanism [I], the government attains budget neutrality as long as problem (18) is feasible (in other words, mechanism [I] is budget neutral by design). Note also that formulation (18) may be infeasible if β is large, for the following reason. If β is large, then the private enterprise must issue a large subsidy to reach the adoption target. However, if the cost of providing a large subsidy exceeds the extra revenue earned by the enterprise (due to increased demand), then the participation constraint (18d) may be violated. Formally, let $\tau(x) = \min\{\frac{K}{((d-p)-x(r-d))} - 1, \bar{\tau}(x)\}$, where $\bar{\tau}(x) = \frac{(\delta-1)(x+1)}{(d-p)-(r-d)x} - 1$ as defined in §2.2. Note that $0 < \tau(x) \leq \bar{\tau}(x)$.

4.1 Indirect Mechanism: Optimal Subsidy z^*

By solving problem (18), we obtain the optimal subsidy z^* in the following proposition.

Proposition 2. *Mechanism [I] is feasible if and only if $\beta \le \tau(x)$. If $\beta \le \tau(x)$, then the optimal subsidy is given by:*

$$z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\}, \tag{19}$$

where $\underline{z} = \beta((d-p) - x(r-d))$, $\hat{z} = \frac{1}{2}(K - ((d-p) - x(r-d)))$ and $\bar{z} = \min\{K, (\delta-1)(x+1)\} - ((d-p) - x(r-d))$. The minimal cost is 0, $C(0,0,z^*) = 0$. Further,

(i) If the enterprise's profit margin K is high, the optimal subsidy z^* strictly increases in x.

(ii) If the enterprise's profit margin K is low, there exists $\tilde{x}$ such that the optimal subsidy z^* strictly decreases on $x < \tilde{x}$ and strictly increases on $x \ge \tilde{x}$.

The intuition behind Proposition 2 is similar to that of Proposition 1. To elaborate, note that as the last mile distance x increases, the mixed mode adoption (before any government intervention) D_m^0 decreases, which generates two competing effects. First, the "adoption effect" (as described in §3.1) continues to play a role here, meaning the requirement for the corresponding mixed mode commuters $D_m(0,0,z)$ becomes less stringent as D_m^0 decreases. Consequently, the private enterprise can afford to reduce the subsidy z. Second, as D_m^0 decreases, the private enterprise has incentive to increase z, so that it can recoup the subsidy cost by increasing demand (i.e, a higher $D_m(0,0,z)$). We shall refer to the latter effect as the "*customer effect*." Note that the customer effect plays a role in mechanism [I] – because the subsidy z is paid by the private enterprise – but is absent in mechanism [D]. Hence, whether the optimal subsidy z^* increases or decreases in x depends on which effect dominates.

To examine which effect dominates, let us consider the case where the enterprise's profit margin K is high. In this case, the enterprise's profit is highly sensitive to changes in D_m^0. Therefore, the private enterprise has a stronger incentive to offer a higher subsidy to boost the customer base $D_m(0,0,z)$, which makes the customer effect dominate the adoption effect. As a result, the subsidy z^* increases in x when K is large. The intuition behind the decrease in z^* when K is small is similar, which we omit for conciseness.

4.2 Comparison: Direct Mechanism [D] versus Indirect Mechanism [I]

By considering Proposition 1 and Proposition 2 presented in §3.1 and §4.1, we can compare the performance of the direct and indirect mechanisms according to each of the six metrics discussed in §2.2. For ease of reference, we shall use superscript D and I to denote quantities obtained at optimal solutions in mechanisms [D] and [I], respectively. Also, in preparation, let $\tilde{\tau}(x) = \frac{\hat{z}^2 - (z^* - \hat{z})^2}{K((d-p)-(r-d)x)}$. The following result compares the performance of the two mechanisms assuming $\beta \le \tau(x)$.

Theorem 1. *Mechanism [I] outperforms mechanism [D] in terms of operating cost ($C^I = 0 < C^D$), mixed mode adoption ($D_m^I \geq D_m^D$), commuter welfare ($W^I > W^D$), ride-hailing platform profit ($\Pi_r^I \geq \Pi_r^D$), and transit agency revenue ($\Pi_p^I \geq \Pi_p^D$). Further,*

(i) if $\beta \leq \tilde{\tau}(x)$, mechanism [I] outperforms mechanism [D] on enterprise profit ($\Pi_s^I \geq \Pi_s^D$),

(ii) if $\beta \in (\tilde{\tau}(x), \tau(x)]$, mechanism [D] outperforms mechanism [I] on enterprise profit ($\Pi_s^I < \Pi_s^D$).

Theorem 1 states that the indirect mechanism [I] dominates the direct mechanism [D] in *all* performance metrics when the adoption target is not too large, $\beta \leq \tilde{\tau}(x)$. However, in the intermediate case where $\tilde{\tau}(x) < \beta \leq \tau(x)$, the direct mechanism [D] will enable the private enterprise to generate a higher profit.

Recall that the main difference between mechanisms [I] and [D] is the funding source for the commuter subsidy. In mechanism [D], the government funds the subsidy by imposing congestion fees on drivers; in mechanism [I], the government simply coordinates the funding from the private sector partner, and is otherwise passive and budget neutral. Theorem 1 highlights the benefits of a public-private partnership over congestion fees: in addition to achieving budget neutrality, the government can both improve commuter welfare (by avoiding congestion fees on drivers) and support the private sector partner (by boosting demand). As a result, all stakeholders are left better off under mechanism [I] compared to [D], assuming the adoption target β is not too large. Moreover, note that mechanism [I] can dominate mechanism [D] despite containing only a single lever (the subsidy z). This result highlights that the gains from mechanism [I] are due to the coordination with the private enterprise.

Also, recall from Proposition 2 that when the adoption target is high $\beta > \tau(x)$, formulation (18) is infeasible, meaning there is no subsidy level z that can satisfy all of the criteria I − VI in §2.2. This observation enables us to make the following summarizing remark:

Remark 2. *If the public transit adoption target is aggressive, $\beta > \tau(x)$, then the government should adopt mechanism [D] over mechanism [I]. If the adoption target is conservative, $\beta \leq \tau(x)$, the government should adopt mechanism [I] over mechanism [D] (even though the private enterprise can earn higher profit under mechanism [D] when β satisfies $\tilde{\tau}(x) < \beta \leq \tau(x)$).*

The above remark has the following implications. When the municipal government has a very tight or no budget, the government should set a conservative adoption target $\beta \leq \tau(x)$. In this case, the government should adopt the indirect mechanism [I], which leaves all parties better off, and critically, does not require increased spending by the government. However, if the local government has ample funding, then it is feasible to set an aggressive transit adoption target level $\beta \geq \tau(x)$, in which case the government should adopt and fund mechanism [D].

5 Alternative Formulation: Maximizing Commuter Welfare Under Budget Neutrality

In §3 and §4, we analyse the optimal incentive schemes under mechanisms [D] and [I] by minimizing operating cost $C(e, s, z)$. In this section, we analyze mechanisms [D] and [I] by considering an alternative objective: maximizing commuter welfare while maintaining budget neutrality. (In Appendix B, we consider a different alternative model formulation for the case when the transit agency has an exogenous given earmark budget B that is reserved for supporting various schemes for increasing public transit ridership. Specifically, we consider two different incentive design problems associated with mechanisms [D] and [I] by maximizing public transit adoption subject to a budget constraint. As shown in Appendix B, we obtain the same structural results under both mechanisms.)

5.1 Alternative Direct Mechanism [D-A]

Recall from Remark 1 in §3.1 that it is not possible to implement the direct mechanism [D] in a budget neutral manner without undermining commuter welfare. However, when a financially constrained municipality finds it too costly to implement mechanism [D], some loss in commuter welfare may be tolerated (especially if there is public support for policies that lower carbon emissions and traffic congestion). This observation motivates us to modify the original incentive design problem (16) by maximizing commuter welfare subject to a budget neutrality constraint (20c). Specifically, the modified problem can be formulated as:

$$\text{Mechanism [D-A]:} \qquad \begin{aligned} &\max_{e,s\geq 0}\ (W(e,s,0) - W^0) && (20a)\\ &\text{s.t. } D_m(e,s,0) - D_m^0 \geq \beta D_m^0 && (20b)\\ &\qquad C(e,s,0) = 0 && (20c)\\ &\qquad \Pi_s(e,s,0) \geq \Pi_s^0, && (20d) \end{aligned}$$

where D_m^0, W^0 and Π_s^0 are the demand for the mixed mode, commuter welfare and profit of the private enterprise in the base model as defined in (3), (4) and (7), respectively. Analogous to Proposition 1, the next result characterizes the optimal incentive scheme under mechanism [D-A].

Proposition 3. *The optimal congestion fee e^* and the optimal subsidy s^* under the alternative direct mechanism [D-A] are given by:*

$$e^* = \frac{\beta(\beta+1)((d-p)-(r-d)x)^2}{(\delta-1)(x+1)}$$

$$s^* = \frac{\beta((d-p)-(r-d)x)((\delta-1)(x+1)-(\beta+1)((d-p)-(r-d)x))}{(\delta-1)(x+1)}$$

Further,

(i) The optimal congestion fee e^* increases in β and decreases in x.

(ii) If δ is large, s^* strictly decreases in x. However, if δ is small, there exists $\tilde{x}$ such that s^* increases on $x < \tilde{x}$ and decreases on $x \geq \tilde{x}$.

(iii) There exists $\tilde{\beta}$ such that s^* increases on $\beta < \tilde{\beta}$ and decreases on $\beta \geq \tilde{\beta}$.

The above results resemble Proposition 1, except that the optimal subsidy s^* under the alternate mechanism [D-A] is no longer always increasing in β. Instead, the optimal subsidy s^* decreases β when the adoption target β is large. This difference is driven by the budget neutrality constraint (20c). To elaborate, consider the case when β is large. In this case, the right hand side of (20b) is large, which requires the mixed mode demand $D_m(e, s, 0)$ to be large. As more mixed mode commuters collect the subsidy, the government must reduce the subsidy s in order to maintain budget neutrality, which leads s^* to decrease in the adoption target β.

Next, similar to Corollary 1, we compare the relative size of the optimal congestion fee and subsidy.

Corollary 3. *There exists $\hat{x}$ such that $e^* \geq s^*$ in mechanism [D-A] if and only if $x \leq \hat{x}$. Further, there exists $\bar{\delta} > 1$ such that $s^* > e^*$ in mechanism [D-A] for all $x \in (\underline{x}, \bar{x})$ if and only if $\delta \geq \bar{\delta}$.*

Corollary 3 is analogous to Corollary 1: for any β, the government should set a higher congestion fee and a lower subsidy when the last mile distance x is small, and vice versa when x is large.

Recall from Corollary 2 in §3.1 that the direct mechanism [D] is costly (i.e., $C(e^*, s^*, 0) > 0$) due to the commuter welfare constraint (16c). Next, we present a counterpart to Corollary 2: to maintain budget neutrality, commuter welfare degradation is unavoidable under the alternative direct mechanism [D-A].

Corollary 4. *The alternative direct mechanism [D-A] will always result in commuter welfare degradation; i.e., $W(e^*, s^*, 0) < W^0$. Further, the reduction in commuter welfare, $W^0 - W(e^*, s^*, 0)$, under mechanism [D-A] is higher when the adoption target β is large and when the last mile distance x is small.*

Corollary 4 reveals that mechanism [D-A] degrades commuter welfare more severely when the adoption target β is large. To elaborate, consider the case when the adoption target β is large. In this case, the corresponding mixed mode demand $D_m(e, s, 0)$ has to be large in order to meet the adoption target. As more mixed mode commuters collect the subsidy, the government must reduce the subsidy in order to maintain budget neutrality, which causes the commuter welfare to deteriorate. Next, we consider why commuter welfare reduction is higher when x is small. In this case, the mixed mode adoption (before any government intervention) D_m^0 as stated in (3) is large, which makes the requirement for the mixed mode adoption $D_m(e, s, 0)$ also large. Similar to the large β case, as more mixed mode commuters receive the subsidy, the government must increase the congestion fee e in order to maintain budget neutrality, which reduces commuter welfare.

5.2 Alternative Indirect Mechanism [I-A]

In the same spirit as §5.1, we now consider a variation of mechanism [I], which we shall refer to as mechanism [I-A], where the objective is to maximize commuter welfare while maintaining budget neutrality. To do so, we modify our original program (18) presented in §4 as follows:

$$\text{Mechanism [I-A]:} \quad \begin{aligned} & \max_{z \geq 0} \; (W(0,0,z) - W^0) && \text{(21a)} \\ & \text{s.t. } D_m(0,0,z) - D_m^0 \geq \beta D_m^0 && \text{(21b)} \\ & \quad C(0,0,z) = 0 && \text{(21c)} \\ & \quad \Pi_s(0,0,z) \geq \Pi_s^0 && \text{(21d)} \\ & \quad z = \arg\max_z \Pi_s(0,0,z) && \text{(21e)} \end{aligned}$$

In parallel with Proposition 2, the following proposition characterizes the optimal subsidy under the alternative indirect mechanism [I-A].

Proposition 4. *The alternative indirect mechanism [I-A] is feasible if and only if $\beta \leq \tau(x)$. When $\beta \leq \tau(x)$, the optimal subsidy is given by*

$$z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\},$$

where $\tau(x)$ is defined in §4 and $\underline{z}, \hat{z}, \bar{z}$ are defined in Proposition 2. The maximal welfare increment is strictly positive so that $W(0,0,z^*) > W^0$.

Note the optimal subsidy z^* under mechanism [I-A] is equal to that under mechanism [I], as presented in Proposition 2. This is because constraints (21b), (21d) and (21e), which are the same as constraints (18b), (18d) and (18e), uniquely determine feasible private subsidy z^* and $C(0,0,z^*) = 0$ by definition in (11). Thus z^* is feasible and thus optimal to both problems (18) and (21). Also, unlike mechanism [D-A] with deteriorating commuter welfare as stated in Corollary 4, Proposition 4 exerts that, when $\beta \leq \tau(x)$, mechanism [I-A] will always enhance commuter welfare.

5.3 Comparison: Two Alternative Mechanisms [D-A] and [I-A]

We now compare the performance of the two alternative mechanisms [D-A] and [I-A]. Using the same approach as presented in §4.2, we can retrieve the relevant performance metrics from (10)–(15) by substituting (e, s, z) with $(e^*, s^*, 0)$ and $(0, 0, z^*)$.

Theorem 2. *The alternative indirect mechanism [I-A] and the alternative direct mechanism [D-A] both attain budget neutrality ($C^I = C^D = 0$). However, alternative indirect mechanism [I-A] outperforms alternative direct mechanism [D-A] in terms of mixed mode adoption ($D_m^I \geq D_m^D$), commuter welfare ($W^I > W^0 > W^D$), ride-hailing platform profit ($\Pi_r^I \geq \Pi_r^D$), and transit agency revenue ($\Pi_p^I \geq \Pi_p^D$). Further,* *(i) if $\beta \leq \tilde{\tau}(x)$, mechanism [I-A] outperforms mechanism [D-A] in terms of enterprise profit ($\Pi_s^I \geq \Pi_s^D$),*

(ii) if $\beta \in (\tilde{\tau}(x), \tau(x)]$, mechanism [D-A] outperforms mechanism [I-A] in terms of enterprise profit ($\Pi_s^I < \Pi_s^D$),

where $\tau(x)$ and $\tilde{\tau}(x)$ are defined in §4 and §2.2, respectively.

Using the above proposition, we summarize our observations in the following remark.

Remark 3. *Suppose in addition to budget neutrality, the government prioritizes commuter welfare. Then the government should adopt the alternative direct mechanism [D-A] when the adoption target satisfies $\beta > \tau(x)$, and adopt the alternative indirect mechanism [I-A] when $\beta \leq \tau(x)$.*

Remark 2 and Remark 3 imply that the dominance of the indirect mechanism [I] is robust, regardless of whether the goal is to maximize commuter welfare or to minimize operating cost. Therefore, when the municipal government is financially constrained, it is advisable for the government to set a conservative transit adoption target $\beta \leq \tau(x)$, and adopt the indirect mechanism [I]. However, if the government has sufficient funding, or if the commuters are willing to accept lower welfare (e.g., in support of a reduction in carbon emissions and traffic congestion), then the government can afford to set an aggressive adoption target $\beta > \tau(x)$ and adopt the direct mechanism [D].

6 Conclusion

Motivated by the increased focus on public transit and urban mobility by municipal governments, we have analyzed two mechanisms for improving public transit ridership. Both mechanisms are intended to address the "last mile gap" by providing subsidies to commuters who adopt a mixed mode of transportation that combines ride-hailing with public transit. The main differences between the two mechanisms are the role of the municipal government and the source of funding for the subsidy. In the direct mechanism [D], the government provides a ride-hailing subsidy to commuters that adopt the mixed mode, and charges a congestion fee to commuters who travel by personal vehicle. The congestion fee is used to offset the cost of the subsidy, and also serves as an additional incentive for transit adoption (by making driving more costly). In the indirect mechanism [I], the government partners with a private enterprise that provides the funding for the subsidy, and does not charge a congestion fee. We present analytical results that characterize the optimal incentives under these two mechanisms, as well as the impact on all involved stakeholders.

Our findings offer several prescriptions for policy makers who are interested in increasing public transit ridership through partnerships with the private sector. First, in the direct mechanism, it's advisable to set large congestion fee and small subsidy when public transit coverage is high (i.e. when the last mile distance is small), and conversely, when the public transit coverage is low, the optimal incentive requires more aggressive subsidies. This result is driven by the opposing effects that subsidies and congestion fees have on commuter welfare and operating cost. We also find that the dependence of the optimal subsidy on the transit coverage level depends on the relative preference between driving and taking public transit. In particular, we find that when commuters strongly prefer driving, jurisdictions with larger last mile distances should set lower subsidies, but this behavior can be reversed if commuters only slightly prefer driving over public transit.

Second, we find that the government cannot fully recover the cost of providing subsidies by collecting congestion fees. Specifically, we show that the direct mechanism cannot be budget neutral unless the government is willing to accept a decrease in commuter welfare. This suggests that attempting to implement commuter subsidies and congestion fees in a budget neutral manner may be ill-advised if the government is sensitive to commuter welfare. However, in the event that the government can obtain subsidy funding from a private enterprise (who benefits from increased foot traffic at the transit station or increased views of an advertisement on a mobile app) in lieu of charging congestion fees (as in the indirect mechanism), we show that public transit ridership can be increased without degrading commuter welfare. However, because the implementation of an indirect mechanism requires the participation of the private enterprise, we find that this indirect mechanism is only viable if the mixed mode adoption target is modest. In other words, only the direct mechanism can enable the government to achieve an ambitious adoption target, and the implementation of the direct mechanism is not budget neutral: it will require extra funding.

Third, although the government has one less "lever" in the indirect mechanism [I] (due to the absence of the congestion fee), it can dominate the direct mechanism [D] if the adoption target is modest. Specifically, when the adoption target is modest, the indirect mechanism can benefit all stakeholders: the government, commuters, the public transit agency, the ride-hailing platform, and the private enterprise. This finding suggests that, for jurisdictions that wish to increase transit ridership, but are severely budget constrained, it may be more fruitful to fund ride-hailing subsidies through partnerships with the private sector than by charging congestion fees. Moreover, this dominance result is robust to an alternate specification of mechanism [I] (namely, when the objective is to maximize commuter welfare instead of enterprise profit).

As an initial attempt to explore different incentive mechanisms for improving public transit ridership, our model has several limitations that we wish to highlight. First, for tractability, we restrict attention to two modes of travel – driving and the mixed mode. Although other means of commuting are less common (such as commuting by ride-hailing), including them as travel modes may affect our results. Second, subsidies for ride-hailing trips may cause commuters to switch from other last-mile travel modes, such as walking, biking, or bussing, which may have consequences not discussed in this paper. Further, to focus on the economic viability of different incentive mechanisms, we have abstracted away from operational issues such as capacity, traffic congestion, and the spatial nature of public transit infrastructure. Considering these features may yield additional insight into the design of incentive mechanisms for improving transit ridership. We shall defer these issues to future research.

We conclude by offering several potential directions for future work. First, it may be fruitful to examine other forms of partnership between the transit agency, ride-hailing platform, and the private enterprise. For example, because the ride-hailing platform benefits from the commuter subsidy (due to increased demand), it may be worthwhile to examine partnerships where the cost of the subsidy is shared between the municipal government and the ride-hailing platform. Second, our work is relevant to other settings where two firms with "complementary capacity" engage in a mutually beneficial partnership. For example, FedEx offers a "SmartPost" delivery service in which FedEx maintains responsibility for the long haul transportation of goods, while the United States Postal Service (USPS) handles the last mile delivery between a USPS center and a customer's home. Our work can serve as a springboard for analyzing partnerships of this nature. Third, we have modeled commuter welfare at an aggregate level; it may also be worthwhile to investigate mechanisms that can improve welfare for *all* commuters (e.g., by reducing traffic congestion for drivers). Lastly, budget neutral incentive mechanisms have received relatively little attention in the operations literature, but may be valuable in other settings where government funds are limited. Our findings suggest that an appropriately designed public-private partnership may be a cost-effective solution for governments that wish to improve adoption of other socially beneficial goods or services (e.g. electric vehicles).

References

Akkaya, Duygu, Kostas Bimpikis, Hau Lee. 2019. Government interventions in promoting sustainable practices in agriculture. *Manufacturing & Service Operations Management* .

Alizamir, Saed, Francis de Véricourt, Peng Sun. 2016. Efficient feed-in-tariff policies for renewable energy technologies. *Operations Research* **64**(1) 52–66.

Alizamir, Saed, Foad Iravani, Hamed Mamani. 2019. An analysis of price vs. revenue protection: Government subsidies in the agriculture industry. *Management Science* **65**(1) 32–49.

American Automobile Association. 2018. Your driving costs: how much are you really paying to drive.

Anandalingam, G. 1987. Asymmetric players and bargaining for profit shares in natural resource development. *Management Science* **33**(8) 1048–1057.

Anderson, Michael L. 2014. Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review* **104**(9) 2763–96.

APTA. 2019a. Apta public transportation fact book. URL `https://www.apta.com/wp-content/uploads/APTA_Fact-Book-2019_FINAL.pdf`.

APTA. 2019b. First last/mile solutions. URL https://www.apta.com/research-technical-resources/mobility-innovation-hub/first-last-mile-solutions/.

APTA. 2019c. Public transportation facts. URL https://www.apta.com/news-publications/public-transportation-facts/.

APTA. 2019d. Transit Agency and Urbanized Area Operating Statistics.

APTA. 2019e. Transit and TNC partnerships. URL https://www.apta.com/research-technical-resources/mobility-innovation-hub/transit-and-tnc-partnerships/.

Arifoglu, Kenan, Christopher S Tang. 2019. A two-sided budget-neutral incentive program for coordinating an influenza vaccine supply chain with endogenous supply and demand under uncertainty. *Forthcoming, Manufacturing & Service Operations Management* .

Aswani, Anil, Zuo-Jun Max Shen, Auyon Siddiq. 2019. Data-driven incentive design in the medicare shared savings program. *Operations Research* **67**(4) 1002–1026.

Badstuber, Nicole. 2018. London's congestion charge is showing its age. URL https://www.citylab.com/transportation/2018/04/londons-congestion-charge-needs-updating/557699/.

Bakshi, Nitin, Noah Gans. 2010. Securing the containerized supply chain: analysis of government incentives for private investment. *Management Science* **56**(2) 219–233.

Bansal, Sangeeta, Shubhashis Gangopadhyay. 2003. Tax/subsidy policies in the presence of environmentally aware consumers. *Journal of Environmental Economics and Management* **45**(2) 333–355.

Brown, Eliot. 2020. The ride-hail utopia that got stuck in traffic. URL https://www.wsj.com/articles/the-ride-hail-utopia-that-got-stuck-in-traffic-11581742802.

Bureau of Transportation Statistics. 2018. Transportation statistics annual report. URL https://www.bts.dot.gov/sites/bts.dot.gov/files/docs/browse-statistical-products-and-data/transportation-statistics-annual-reports/Preliminary-TSAR-Full-2018-a.pdf.

Burkhauser, Richard V, Timothy M Smeeding. 1994. Social security reform: A budget neutral approach to reducing older women's disproportional risk of poverty. *Available at SSRN 1824592* .

Chemama, Jonathan, Maxime C Cohen, Ruben Lobel, Georgia Perakis. 2019. Consumer subsidies with a strategic supplier: Commitment vs. flexibility. *Management Science* **65**(2) 681–713.

Chen, Linxi, Hai Yang. 2012. Managing congestion and emissions in road networks with tolls and rebates. *Transportation Research Part B: Methodological* **46**(8) 933–948.

Cho, Soo-Haeng, Xin Fang, Sridhar Tayur, Ying Xu. 2019. Combating child labor: Incentives and information disclosure in global supply chains. *Manufacturing & Service Operations Management* **21**(3) 692–711.

CMS. 2018. Budget neutrality policies for section 1115(a) medicaid demonstration projects. URL https://www.medicaid.gov/sites/default/files/federal-policy-guidance/downloads/smd18009.pdf.

Cohen, Morris A, Shiliang Cui, Ricardo Ernst, Arnd Huchzermeier, Panos Kouvelis, Hau L Lee, Hirofumi Matsuo, Marc Steuber, Andy A Tsay. 2018. Om forum—benchmarking global production sourcing decisions: Where and why firms offshore and reshore. *Manufacturing and Service Operations Management* **20**(3) 389–402.

Cohen, Morris A, Hau L Lee. 2020. Designing the right global supply chain network. *Manufacturing & Service Operations Management* .

Cole, Matt. 2017. Mobility as a service: Putting transit front and center of the conversation. *Cubic White Paper* .

Condor, Bob. 2020. Get me to the game on time. URL `https://www.nhl.com/kraken/news/nhl-seattle-announces-monorail-partnership-and-free-public-transit/c-315511164`.

Correia, Isabel, Emmanuel Farhi, Juan Pablo Nicolini, Pedro Teles. 2013. Unconventional fiscal policy at the zero bound. *American Economic Review* **103**(4) 1172–1211.

Corselli, Andrew. 2020. Cubic reveals public transit loyalty program. URL `https://www.railwayage.com/passenger/cubic-reveals-public-transit-loyalty-program/`.

Croci, Edoardo. 2016. Urban road pricing: a comparative study on the experiences of london, stockholm and milan. *Transportation Research Procedia* **14** 253–262.

Currie, Graham. 2018. Lies, damned lies, avs, shared mobility, and urban transit futures. *Journal of Public Transportation* **21**(1) 3.

Development Asia. 2018. The case for electronic road pricing. URL `https://development.asia/case-study/case-electronic-road-pricing`.

D'Acunto, Francesco, Daniel Hoang, Michael Weber. 2016. The effect of unconventional fiscal policy on consumption expenditure. Tech. rep., National Bureau of Economic Research.

Edwards, Tom. 2018. What's to blame for a fall in passenger numbers? URL `https://www.bbc.com/news/uk-england-london-42622891`.

Fare Estimater. 2019. Los Angeles Lyft Prices. URL `https://estimatefares.com/rates/los-angeles`.

Federal Highway Administration. 2017. National household travel survey.

Federal Transit Administration. 2017. NTD 2017 Data Tables.

Federal Transit Administration. 2020. 2019 annual database revenue sources. URL `https://www.transit.dot.gov/ntd/data-product/2019-annual-database-revenue-sources`.

Gagnepain, Philippe, Marc Ivaldi. 2002. Incentive regulatory policies: the case of public transit systems in france. *RAND Journal of Economics* 605–629.

Goodall, Warwick, Tiffany Dovey, Justine Bornstein, Brett Bonthron. 2017. The rise of mobility as a service. *Deloitte Rev* **20** 112–129.

Goulder, Lawrence H. 1995. Environmental taxation and the double dividend: a reader's guide. *International tax and public finance* **2**(2) 157–183.

Guan, Peiqiu, Jing Zhang, Vineet M Payyappalli, Jun Zhuang. 2018. Modeling and validating public–private partnerships in disaster management. *Decision Analysis* **15**(2) 55–71.

Guan, Peiqiu, Jun Zhuang. 2015. Modeling public–private partnerships in disaster management via centralized and decentralized models. *Decision Analysis* **12**(4) 173–189.

Guo, Pengfei, Robin Lindsey, Zhe George Zhang. 2014. On the downs–thomson paradox in a self-financing two-tier queuing system. *Manufacturing & Service Operations Management* **16**(2) 315–322.

Guo, Pengfei, Christopher S Tang, Yulan Wang, Ming Zhao. 2019. The impact of reimbursement policy on social welfare, revisit rate, and waiting time in a public healthcare system: Fee-for-service versus bundled payment. *Manufacturing & Service Operations Management* **21**(1) 154–170.

Guo, Xiaolei, Hai Yang. 2010. Pareto-improving congestion pricing and revenue refunding with multiple user classes. *Transportation Research Part B: Methodological* **44**(8-9) 972–982.

Gupta, Diwakar, Mili Mehrotra. 2015. Bundled payments for healthcare services: Proposer selection and information sharing. *Operations Research* **63**(4) 772–788.

Gur, Yonatan, Lijian Lu, Gabriel Y Weintraub. 2017. Framework agreements in procurement: An auction model and design recommendations. *Manufacturing & Service Operations Management* **19**(4) 586–603.

Hansson, Lisa. 2010. Solving procurement problems in public transport: Examining multi-principal roles in relation to effective control mechanisms. *Research in Transportation Economics* **29**(1) 124–132.

Hodges, Tina, et al. 2010. Public transportation's role in responding to climate change. *United States. Federal Transit Administration* .

Homonoff, Tatiana A. 2018. Can small incentives have large effects? the impact of taxes versus bonuses on disposable bag use. *American Economic Journal: Economic Policy* **10**(4) 177–210.

Jou, Rong-Chang, Yu-Chiun Chiou, Ke-Hong Chen, Hao-I Tan. 2012. Freeway drivers' willingness-to-pay for a distance-based toll rate. *Transportation Research Part A: Policy and Practice* **46**(3) 549–559.

Kamp, Jon. 2020. Cities Offer Free Buses in Bid to Boost Flagging Ridership. URL `https://www.wsj.com/articles/cities-waive-bus-fares-in-bid-to-boost-ridership-11579007895`.

Kang, Chao-Chung, Tsun-Siou Lee, Szu-Chi Huang. 2013. Royalty bargaining in public–private partnership projects: insights from a theoretic three-stage game auction model. *Transportation research part E: logistics and transportation review* **59** 1–14.

Kirk, Robert S, William J Mallett. 2020. Highway and public transit funding issues. *Congressional Research Service.* URL `https://crsreports.congress.gov/product/pdf/download/IF/IF10495/IF10495.pdf/`.

LA Metro. 2016. Metro first/last mile.

Lee, Donald KK, Stefanos A Zenios. 2012. An evidence-based incentive system for medicare's end-stage renal disease program. *Management science* **58**(6) 1092–1105.

Leonard, Christian. 2019. Study: Irvine workers spend nearly $10,000 a year in commute. URL https://www.nbclosangeles.com/news/local/irvine-workers-spend-10000-commuting-565069612.html.

Levi, Retsef, Georgia Perakis, Gonzalo Romero. 2017. On the effectiveness of uniform subsidies in increasing market consumption. *Management Science* **63**(1) 40–57.

Linton, Joe. 2016. New map shows metro's 20,000 parking spaces, mostly free. URL https://la.streetsblog.org/2016/09/23/new-map-shows-metros-20000-parking-spaces-mostly-free/.

Lodi, Andrea, Enrico Malaguti, Nicolás E Stier-Moses, Tommaso Bonino. 2016. Design and control of public-service contracts and an application to public transportation systems. *Management Science* **62**(4) 1165–1187.

Los Angeles County Metropolitan Transportation Authority. 2019. La metro launches partnership with via to provide on-demand service to three busy transit stations. URL https://www.metro.net/news/simple_pr/la-metro-launches-partnership-provide-demand-servi/.

Ma, Guangrui, Michael K Lim, Ho-Yin Mak, Zhixi Wan. 2019. Promoting clean technology adoption: To subsidize products or service infrastructure? *Service Science* **11**(2) 75–95.

Mallett, William J. 2018. Trends in public transportation ridership: implications for federal policy. *Washington, DC: Congressional Research Service* .

Murray, Brian, Nicholas Rivers. 2015. British columbia's revenue-neutral carbon tax: A review of the latest "grand experiment" in environmental policy. *Energy Policy* **86** 674–683.

Nie, Yu Marco, Yang Liu. 2010. Existence of self-financing and pareto-improving congestion pricing: Impact of value of time distribution. *Transportation Research Part A: Policy and Practice* **44**(1) 39–51.

Rastogi, Rajat, KV Krishna Rao. 2003. Travel characteristics of commuters accessing transit: Case study. *Journal of Transportation Engineering* **129**(6) 684–694.

Saban, Daniela, Gabriel Y Weintraub. 2019. Procurement mechanisms for assortments of differentiated products. *Working paper, Stanford University, California* .

Schwieterman, Joseph, Mallory Livingston. 2018. Partners in transit: A review of partnerships between transportation network companies and public agencies in the united states .

So, Kut C, Christopher S Tang. 2000. Modeling the impact of an outcome-oriented reimbursement policy on clinic, patients, and pharmaceutical firms. *Management Science* **46**(7) 875–892.

Taylor, Terry A, Wenqiang Xiao. 2014. Subsidizing the distribution channel: Donor funding to improve the availability of malaria drugs. *Management Science* **60**(10) 2461–2477.

The Economist. 2019. Public transport is in decline in many wealthy cities. URL https://www.economist.com/international/2018/06/21/public-transport-is-in-decline-in-many-wealthy-cities.

Tomer, Adie, Elizabeth Kneebone, Robert Puentes, Alan Berube. 2011. Missed opportunity: Transit and jobs in metropolitan america. URL https://www.brookings.edu/wp-content/uploads/2016/06/0512_jobs_transit.pdf.

Transport for London. 2019. Congestion charge (official). URL https://tfl.gov.uk/modes/driving/congestion-charge.

Uber. 2020. Make your train with uber transit. URL https://www.uber.com/newsroom/make-my-train/.

US Census Bureau. 2018. Means of transportation to work by selected characteristics. *American Community Survey* .

Van Ommeren, Jos, Gerard J Van den Berg, Cees Gorter. 2000. Estimating the marginal willingness to pay for commuting. *Journal of regional science* **40**(3) 541–563.

Wang, Hua, Xiaoning Zhang. 2016. Joint implementation of tradable credit and road pricing in public-private partnership networks considering mixed equilibrium behaviors. *Transportation Research Part E: Logistics and Transportation Review* **94** 158–170.

Weimer, David L, Aidan R Vining. 2017. *Policy analysis: Concepts and practice*. Routledge.

Xiao, Feng, HM Zhang. 2014. Pareto-improving and self-sustainable pricing for the morning commute with nonidentical commuters. *Transportation Science* **48**(2) 159–169.

Xiao, Ping, Ruli Xiao, Yitian Liang, Xinlei Chen, Wei Lu. 2019. The effects of a government's subsidy program: Accessibility beyond affordability. *Management Science* .

Yang, Nan, Yong Long Lim. 2018. Temporary incentives change daily routines: Evidence from a field experiment on singapore's subways. *Management Science* **64**(7) 3365–3379.

Yu, Jiayi Joey, Christopher S Tang, Zuo-Jun Max Shen. 2018. Improving consumer welfare and manufacturer profit via government subsidy programs: Subsidizing consumers or manufacturers? *Manufacturing & Service Operations Management* **20**(4) 752–766.

Zhu, Jing, Yingling Fan. 2018. Daily travel behavior and emotional well-being: Effects of trip mode, duration, purpose, and companionship. *Transportation Research Part A: Policy and Practice* **118** 360–373.

A General Hybrid Mechanism: Optimal Incentive (e^*, s^*, z^*)

In the previous sections, we find that mechanisms [D] and [I] capture two types of incentives that have already been implemented in practice: the funding of transit subsidies through congestion fees and partnerships with a private enterprise, respectively. By comparing both mechanisms, we found that mechanism [D] is costly to implement, but is capable of achieving higher public transit adoption target. Conversely, mechanism [I] is budget neutral, but only viable for small β. This raises a natural question: Is it possible to maintain budget neutrality while achieving a higher adoption target by combining both mechanisms?

In this section, we propose a general mechanism – denoted by [G] – that combines mechanisms [D] and [I]. In mechanism [G], all three levers, namely, the congestion fees e, the government subsidy s and the private subsidy z, are available. In other words, we allow $e \geq 0$, $s \geq 0$ and $z \geq 0$ in the general formulation presented in §2.2. Figure 4 depicts the relationships among the various stakeholders under the general mechanism [G].

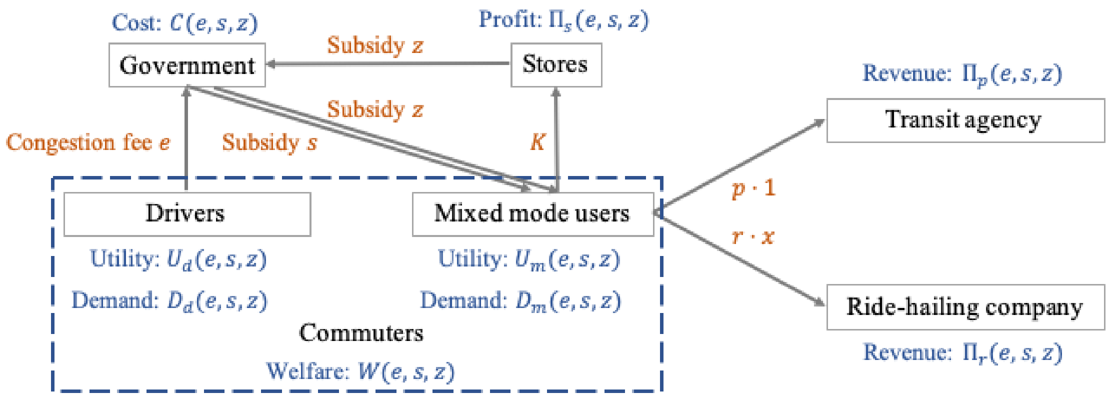

Figure 4: Strategic interactions among stakeholders under the general mechanism [G].

Based on criteria I – VI, the optimal incentive (e^*, s^*, z^*) under mechanism [G] is thus given by the solution to the following optimization problem. Note that similar to mechanism [I], because mechanism [G] requires the participation of the private enterprise, the subsidy z is set to be profit maximizing.

$$\text{Mechanism [G]:} \quad \begin{aligned} \min_{e,s,z\geq 0} \quad & [C(e,s,z)]^+ && \text{(22a)} \\ \text{s.t.} \quad & D_m(e,s,z) - D_m^0 \geq \beta D_m^0 && \text{(22b)} \\ & W(e,s,z) \geq W^0 && \text{(22c)} \\ & \Pi_s(e,s,z) \geq \Pi_s^0 && \text{(22d)} \\ & z = \arg\max_z \Pi_s(e,s,z). && \text{(22e)} \end{aligned}$$

We now characterize the optimal incentive (e^*, s^*, z^*) attained at a solution to (22). For ease of exposition, let $\gamma_1 \equiv (\delta-1)(x+1)$, $\gamma_2 \equiv (d-p)-x(r-d)$, $\breve{x} \equiv \frac{(d-p)-(\delta-1)+2K}{(r-d)+(\delta-1)}$, and $\beta_1(x) \equiv \frac{2\gamma_1-3\gamma_2-\sqrt{(2\gamma_1-\gamma_2)^2-8\gamma_1 K}}{2\gamma_2} < \bar{\tau}(x)$, $\bar{\beta}(x) \equiv \frac{2K}{\gamma_2}$, where $\bar{\beta}(x) < \bar{\tau}(x)$ if and only if $x > \breve{x}$.

Proposition 5. *(i) If the last mile distance is small, $x \leq \breve{x}$, then for any $\beta \in (0, \bar{\tau}(x))$, there exists an optimal incentive program (e^*, s^*, z^*) that incurs no operating cost for the government, $[C(e^*, s^*, z^*)]^+ = 0$. Further,*

(a) if $0 \leq \beta \leq \tau(x)$, then $e^* \geq 0$, $s^* \geq 0$, and $z^* > 0$;

(b) if $\tau(x) < \beta < \bar{\tau}(x)$, then $e^* > 0$, $s^* \geq 0$, and $z^* > 0$.

(ii) Suppose the last mile distance is large, $x > \breve{x}$. Then

(a) if $0 \leq \beta \leq \bar{\beta}(x)$, there exists an optimal incentive program (e^*, s^*, z^*) that incurs no operating cost for the government, $[C(e^*, s^*, z^*)]^+ = 0$. Further,

(1) *if* $0 \le \beta \le \tau(x)$*, then* $e^* \ge 0$*,* $s^* \ge 0$*, and* $z^* > 0$*;*

(2) *if* $\tau(x) < \beta \le \beta_1(x)$*, then* $e^* > 0$*,* $s^* \ge 0$ *and* $z^* > 0$*;*

(3) *if* $\beta_1(x) < \beta \le \bar{\beta}(x)$*, then* $e^* > 0$*,* $s^* > 0$*, and* $z^* > 0$*.*

(b) *if* $\bar{\beta}(x) < \beta < \bar{\tau}(x)$*, then the optimal incentive program* (e^*, s^*, z^*) *incurs a positive operating cost for the government,* $[C(e^*, s^*, z^*)]^+ > 0$*. Further,* $e^* > 0$*,* $s^* > 0$*, and* $z^* > 0$*.*

We offer a few remarks on Proposition 5. First, under mechanism [G], when the adoption target is modest, $\beta \le \bar{\beta}(x)$, budget neutrality is attainable; when the adoption target is aggressive, $\beta > \bar{\beta}(x)$, the government has to bear a positive operating cost. Therefore, analogous to Remark 2, we can conclude that when adopting mechanism [G], a financially strained government should conservatively set the adoption target β to attain budget neutrality, while a sufficiently funded government can set a more aggressive target and further improve public transit adoption. Second, because $\tau(x) < \bar{\beta}(x)$ for any fixed x, it follows that mechanism [G] allows the government to remain budget neutral while achieving higher transit adoption target β than mechanism [I] and leaving all stakeholders (weakly) better off. Third, because the optimal private subsidy is always strictly positive, namely, $z^* > 0$ for all $\beta > 0$, we conclude that it is essential for the municipal government to partner with the private sector if the objective is to minimize operating cost. In addition, as β increases, to achieve more a ambitious adoption target, more interventions should be utilized, that is, $e^* > 0$ and $s^* > 0$.

Using the optimal solutions to mechanism [G], we are able to compare the optimal cost with mechanisms [D] and [I], which is summarized in Corollary 5.

Corollary 5. *For any adoption target* $\beta \in (0, \bar{\tau}(x))$*, the government should always adopt mechanism [G] over mechanisms [D] and [I]. Further, the minimal operating cost* $[C(e^*, s^*, z^*)]^+$ *is higher when* β *is large.*

The above corollary is driven by the fact that the optimal government cost under mechanism [G] is always less than (or equal to) mechanisms [D] and [I]. Intuitively, this occurs because the government can fund the commuter subsidy from both congestion fees and the private enterprise, instead of from just a single source. Specifically, when the adoption target is conservative, $\beta \le \tau(x)$, both mechanisms [G] and [I] are budget neutral; when the target is ambitious, $\tau(x) < \beta < \bar{\tau}(x)$, the optimal cost under mechanism [G] is strictly less than that of mechanism [D]. The reason why the minimal operating cost under mechanism [G] increases in β is similar to intuition for mechanism [D], which we omit for conciseness.

To summarize, we find that when combining both mechanisms, mechanism [G] attains budget neutrality when β is modest, and incurs a strictly positive cost when β is aggressive. Specifically, in the budget neutral regime, mechanism [G] can satisfy a higher adoption target β than mechanism [I], and in the positive cost regime, mechanism [G] is less costly than mechanism [D]. To the best of our knowledge, there is no incentive mechanism that exists in practice that resembles mechanism [G]. Therefore, it may be worthwhile for municipal governments to investigate incentive programs that combine multiple sources of funding for last-mile subsidies (e.g., both congestion fees and private sector partnerships).

B Alternative Formulation: Maximizing Public Transit Adoption as Objective Function

In addition to minimizing operating cost and maximizing commuter welfare, some municipal governments may have an earmarked budget for incentivizing public transit usage. To address this, in this section, we consider the incentive design problem of maximizing public transit adoption with a strictly positive budget $B > 0$, under both mechanisms [D] and [I].

B.1 Mechanism [D-O] with Transit Adoption Objective

In this section, we consider a variation of mechanism [D] where the objective is to maximize public transit adoption with a budget constraint, which we refer to as mechanism [D-O]. We then modify the original incentive design problem (16) as follows:

$$\text{Mechanism [D-O]:} \quad \begin{aligned} \min_{e,s\geq 0} \ & D_m(e,s,0) - D_m^0 && \text{(23a)} \\ \text{s.t. } & C(e,s,0) \leq B && \text{(23b)} \\ & W(e,s,0) \geq W^0 && \text{(23c)} \\ & \Pi_s(e,s,0) \geq \Pi_s^0. && \text{(23d)} \end{aligned}$$

Note that when the budget B is very large, the government can induce all commuters to adopt the mixed mode by simply setting a high subsidy and not charging a congestion fee. To avoid this degenerate case, we assume $B < \bar{B}$, where $\bar{B} \equiv (\delta-1)(x+1) - ((d-p) - x(r-d))$. Then by solving problem (23), we obtain the optimal incentives e^* and s^* under Mechanism [D-O], given in Proposition 6 below. In preparation, let $\tilde{B} = \frac{((\delta-1)(x+1)-((d-p)-x(r-d)))^2}{2(\delta-1)(x+1)}$. It is straightforward to verify that $\tilde{B} < \bar{B}$.

Proposition 6. *For fixed x, the optimal incentives (e^*, s^*) under mechanism [D-O] depends on the size of the budget B. In particular,*

(i) When the budget is small, $B \leq \tilde{B}$, there exists a unique solution

$$\begin{aligned} e^* &= \frac{\sqrt{2B}((d-p)-x(r-d))}{\sqrt{(\delta-1)(x+1)}} + B, \\ s^* &= B\left(\frac{\sqrt{2}((\delta-1)(x+1)-((d-p)-x(r-d)))}{\sqrt{B(\delta-1)(x+1)}} - 1\right). \end{aligned} \tag{24}$$

Upon implementation of the incentive program $(e^*, s^*, 0)$, mixed mode demand will increase to $D_m(e^*, s^*, 0) = \frac{\sqrt{2B(\delta-1)(x+1)}+(d-p)-x(r-d)}{(\delta-1)(x+1)}$.

(ii) When the budget is large, $B > \tilde{B}$, the optimal subsidy is in the interval $s^* \in [\tilde{B}, B]$. Further, for every $s^* \in [\tilde{B}, B]$, the corresponding optimal congestion fee satisfies $e^* \geq \bar{B} - s^*$, and all commuters take the mixed mode, $D_m(e^*, s^*, 0) = 1$.

Note that the mixed mode demand under the optimal incentive program is non-decreasing in the government budget B. This result is intuitive - as the government's budget increases, so does the size of the subsidy, which increases transit adoption. Because the number of commuters are limited, when the budget B is sufficiently large, it enables the government to provide a large enough subsidy such that all the commuters choose public transit, in which case the mixed mode demand is constant at $D_m(e^*, s^*, 0) = 1$.

B.2 Mechanism [I-O] with Transit Adoption Objective

Next, we consider a variant of mechanism [I] with the objective of maximizing public transit adoption, which we call mechanism [I-O]. The optimal incentive z^* under mechanism [I-O] is given by the solution to the following optimization problem:

$$\text{Mechanism [I-O]:} \quad \begin{aligned} &\max_{z \geq 0} D_m(0,0,z) - D_m^0 && \text{(25a)} \\ &\text{s.t. } C(0,0,z) \leq B && \text{(25b)} \\ &\quad W(0,0,z) \geq W^0 && \text{(25c)} \\ &\quad \Pi_s(0,0,z) \geq \Pi_s^0 && \text{(25d)} \\ &\quad z = \arg\max_z \Pi_s(0,0,z) && \text{(25e)} \end{aligned}$$

By solving problem (25), we obtain the optimal subsidy z^* under mechanism [I] as follows.

Proposition 7. *The optimal subsidy z^* under mechanism [I-O] is given by*

$$z^* = \min\left\{\hat{z}, (\delta-1)(x+1) - ((d-p) - x(r-d))\right\},$$

where $\hat{z} = \frac{1}{2}(K - ((d-p) - x(r-d)))$ as defined in Proposition (2)*. The resulting mixed mode demand is $D_m(0,0,z^*) = \frac{\min\{\frac{1}{2}(K+((d-p)-x(r-d))),(\delta-1)(x+1)\}}{(\delta-1)(x+1)}$.*

Proposition 7 reveals that the resulting transit demand under the optimal private subsidy program z^* is non-decreasing in the partner enterprise's unit profit K. To see why, note that when the enterprise's profit margin K is high, the enterprise's profit is highly sensitive to changes in D_m^0, so the private enterprise has a stronger incentive to boost the customer base $D_m(0,0,z)$. As a result, the optimal transit demand $D_m(0,0,z^*)$ is non-decreasing in K. However, when the profit margin K is sufficiently large, the enterprise sets a large enough subsidy z such that all commuters adopt the mixed mode, in which case the mixed mode demand is constant, $D_m(0,0,z^*) = 1$.

To summarize, the optimal mixed mode adoption under mechanism [D-O] (where the private enterprise is passive) depends on the government's budget B, and the optimal mixed mode adoption under mechanism [I-O] (where the government is passive) depends on the enterprise's profit margin K. Note that due to the additional budget parameter B, it is difficult to compare the relative effectiveness of mechanism [D-O] and mechanism [I-O] with respect to the government's operating cost and commuter welfare.

C Commuter Heterogeneity Captured by Last Mile Length x

In the main mode, we capture commuter heterogeneity by assuming that commuters' unit valuation for the mixed mode V is uniformly distributed between 0 and 1, for any fixed last mile distance x. In this section, we assume V is deterministic and identical for all commuters, and instead the last mile x distance is uniformly distributed. Specifically, we assume commuters are uniformly located in the suburban area of length σ with density 1[20], so the mass of commuters is σ and their last-mile distance is $x \sim U[0,\sigma]$. In this case, we also

[20] Our results also persist when the commuter density is generalized to be a constant λ. We assume the commuter density to be 1 for ease of exposition and in parallel with the original model.

aim to find the optimal incentives under both mechanisms [D] and [I] based on criteria I – VI as defined in §2. The related metrics are defined as follows.

Note that in this case, commuter utilities for the mixed mode and driving are the same as defined in (8). Using (8), we have that under the incentive (e, s, z), a commuter will adopt the mixed mode if and only if $U_m(e, s, z) \geq U_d(e, s, z)$, or equivalently, if her last mile distance x satisfies:

$$x \leq x_1(e, s, z) \equiv \min\left\{\sigma, \frac{(d-p)-(\delta-1)V+e+s+z}{(r-d)+(\delta-1)V}\right\}. \tag{26}$$

The demand for each mode is then derived as:

$$D_d(e, s, z) = \left[\sigma - \frac{(d-p)-(\delta-1)V+e+s+z}{(r-d)+(\delta-1)V}\right]^+, \tag{27a}$$

$$D_m(e, s, z) = \sigma - \left[\sigma - \frac{(d-p)-(\delta-1)V+e+s+z}{(r-d)+(\delta-1)V}\right]^+. \tag{27b}$$

Let $D_d^0 = D_d(0,0,0)$ and $D_m^0 = D_m(0,0,0)$. To exclude the degenerate cases where all commuters adopt the same travel mode, we assume that $V \in (\underline{V}, \bar{V})$, so that $D_m^0 > 0$ and $D_d^0 > 0$, where $\underline{V} = \max\{0, \frac{d-p-(r-d)\sigma}{(\delta-1)(\sigma+1)}\}$ and $\bar{V} = \frac{d-p}{\delta-1}$. Analogously, based on the transit adoption criterion II, we also assume throughout that the target β is restricted to the non-degenerate case where $D_d(e, s, z) > 0$ and $D_m(e, s, z) > 0$. To enforce the preceding inequalities, it suffices to assume that $\beta < \bar{\tau}(x) \equiv \frac{\sigma((r-d)+(\delta-1)V)}{(d-p)-(\delta-1)V} - 1$.

Next, we define the performance metrics to evaluate incentive (e, s, z) in the same fashion as §2.2. The total welfare accrued to all commuters is:

$$W(e, s, z) = \int_0^{x_1(e,s,z)} U_m(e, s, z)dV + \int_{x_1(e,s,z)}^{\sigma} U_d(e, s, z)dV, \tag{28}$$

and the operating cost of the government, public transit agency's revenue, ride-hailing platform's revenue and the total profit of the partner enterprise are the same as defined in (11), (13), (14), and (15). We denote the various metrics in the absence of any incentives (i.e., when $(e, s, z) = 0$) in the same way as in §2, namely by using superscript 0: $W(0,0,0) = W^0$, $\Pi_p(0,0,0) = \Pi_p^0$, $\Pi_r(0,0,0) = \Pi_r^0$, and $\Pi_s(0,0,0) = \Pi_s^0$.

C.1 Mechanism [D-H] with Heterogeneous x

In this section, we investigate the optimal congestion fees and government subsidies under mechanism [D] when assuming last mile distance $x \sim U[0, \sigma]$, which we shall refer to as mechanism [D-H]. Based on the same criteria given in I – VI, the optimal incentives e^* and s^* are given by the solution to the following problem:

$$\min_{e,s\geq 0} \; [C(e, s, 0)]^+ \tag{29a}$$

$$\text{s.t.} \; D_m(e, s, 0) - D_m^0 \geq \beta D_m^0 \tag{29b}$$

$$\text{Mechanism [D-H]:} \quad W(e, s, 0) \geq W^0 \tag{29c}$$

$$\Pi_s(e, s, 0) \geq \Pi_s^0. \tag{29d}$$

Proposition 8. *The optimal congestion fee e^* and the optimal subsidy s^* under mechanism [D-H] are given* *by:*

$$\begin{aligned} e^* &= \frac{\beta(\beta+2)((d-p)-(\delta-1)V)^2}{2\sigma((r-d)+(\delta-1)V)}, \\ s^* &= \beta((d-p)-(\delta-1)V) - \frac{\beta(\beta+2)((d-p)-(\delta-1)V)^2}{2\sigma((r-d)+(\delta-1)V)}. \end{aligned} \tag{30}$$

Further, the government incurs positive operating cost: $C(e^*, s^*, 0) = \frac{\beta^2((d-p)-(\delta-1)V)^2}{2((r-d)+(\delta-1)V)} > 0$.

The above proposition resembles the results given in §3.1. Similar to mechanism [D], mechanism [D-H] enables the government to achieve the public transit adoption target β, but it takes a positive cost to implement, and the minimal operating cost $C(e^*, s^*, 0)$ increases in the adoption target β.

C.2 Mechanism [I-H] with Heterogeneous x

Analogous to §C.1, when assuming $x \sim U[0, \sigma]$, we denote mechanism [I] as mechanism [I-H]. Similar to the previous sections, we assume the participating enterprise has sufficient profit margin $K > \underline{K} = (d-p) - (\delta-1)V$. The optimal incentive z^* under mechanism [I-H] is then given by the solution of the following optimization problem:

$$\max_{z \geq 0} [C(0,0,z)]^+ \tag{31a}$$

$$\text{s.t. } D_m(0,0,z) - D_m^0 \geq \beta D_m^0 \tag{31b}$$

$$\text{Mechanism [I-H]:} \quad W(0,0,z) \geq W^0 \tag{31c}$$

$$\Pi_s(0,0,z) \geq \Pi_s^0 \tag{31d}$$

$$z = \arg\max_z \Pi_s(0,0,z) \tag{31e}$$

Similar to formulation (18), the optimization problem (31) may be infeasible if β is large. The intuition is presented in §4, which we omit for conciseness. Define $\tau(x) = \min\{\frac{K}{(d-p)-(\delta-1)V} - 1, \bar{\tau}(x)\}$, where $\bar{\tau}(x) = \frac{\sigma((r-d)+(\delta-1)V)}{(d-p)-(\delta-1)V} - 1$. Note that $0 < \tau(x) \leq \bar{\tau}(x)$.

Proposition 9. *Mechanism [I-H] is feasible if and only if $\beta \leq \tau(x)$. When $\beta \leq \tau(x)$, then the optimal subsidy z^* is given by:*

$$z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\}, \tag{32}$$

where $\underline{z} = \beta((d-p)-(\delta-1)V)$, $\hat{z} = \frac{1}{2}(K - ((d-p)-(\delta-1)V))$ *and* $\bar{z} = \min\{K, \sigma((r-d)+(\delta-1)V)\} - ((d-p)-(\delta-1)V)$. *Further, the government incurs no operating cost:* $C(0,0,z^*) = 0$.

Analogous to §4.2, using the optimal incentives under mechanisms [D-H] and [I-H], we can compare the the performance of these two mechanisms. We use superscript D and I to denote quantities obtained at optimal solutions in mechanisms [D-H] and [I-H], respectively. In preparation, define $\tilde{\tau}(x) = \frac{\hat{z}^2 - (\hat{z}-z^*)^2}{K((d-p)-(\delta-1)V)}$.

Theorem 3. *The indirect mechanism [I-H] outperforms the alternative direct mechanism [D-H] on operating cost ($C^I = 0 < C^D$), mixed mode adoption ($D_m^I \geq D_m^D$), commuter welfare ($W^I > W^D$), ride-hailing platform profit ($\Pi_r^I \geq \Pi_r^D$), and transit agency revenue ($\Pi_p^I \geq \Pi_p^D$). Further,* *(i) if $\beta \le \tilde{\tau}(x)$, mechanism [I-H] outperforms mechanism [D-H] on enterprise profit ($\Pi_s^I \ge \Pi_s^D$),*

(ii) if $\beta \in (\tilde{\tau}(x), \tau(x)]$, mechanism [D-H] outperforms mechanism [I-H] on enterprise profit ($\Pi_s^I < \Pi_s^D$),

Comparing Theorem 3 to Theorem 1, we find that the dominance of the indirect mechanism is robust, regardless of whether the commuters heterogeneity is captured by random unit utility V or random last mile distance x. We draw the following conclusions in parallel with §4.2: When the adoption target is conservative, $\beta \le \tau(x)$, the government should adopt the indirect mechanism [I-H]; when the adoption target is aggressive, $\beta > \tau(x)$, the government should adopt the direct mechanism [D-H].

D Proofs

Lemma 1. *In mechanism [D], the following inequality holds for all (e, s) such that $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$, or equivalently, $e+s \le (\delta-1)(x+1) - ((d-p)-(r-d)x)$:*

$$[C(e,s,0)]^+ \ge C(e,s,0) \ge \frac{(\beta((d-p)-(r-d)x))^2}{2(\delta-1)(x+1)}.$$

Proof Note that when requiring $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$, constraint (16b) implies

$$D_m(e,s,0) - D_m^0 = \frac{e+s}{(\delta-1)(x+1)} \ge \beta D_m^0 = \beta\frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}.$$

Because $\delta > 1$ and $x > 0$, it follows that

$$e+s \ge \beta((d-p)-(r-d)x), \tag{33}$$

which holds with equality only if constraint (16b) is binding. Next, using (11), the operational cost is given by

$$C(e,s,0) = sD_m(e,s,0) - eD_d(e,s,0) = \frac{e((d-p)-(r-d)x-(\delta-1)(x+1)) + s((d-p)-(r-d)x)}{(\delta-1)(x+1)} + \frac{(e+s)^2}{(\delta-1)(x+1)}.$$

Further, replacing (e,s,z) with $(e,s,0)$ in (8) and (12), we obtain $U_d(e,s,0)$, $U_m(e,s,0)$ and $W(e,s,0)$. Then the welfare *increment* is given by

$$\begin{aligned} W(e,s,0) - W^0 &= \left[\int_0^{v(e,s,0)} U_m(e,s,0)dV + \int_{v(e,s,0)}^1 U_d(e,s,0)dV\right] - \left[\int_0^{v^0} U_m^0 dV + \int_v^1 U_d^0 dV\right] &(34)\\ &= \frac{e((d-p)-(r-d)x-(\delta-1)(x+1)) + s((d-p)-(r-d)x)}{(\delta-1)(x+1)} + \frac{(e+s)^2}{2(\delta-1)(x+1)}. &(35)\end{aligned}$$

It follows that for all (e,s) such that $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$, it holds that

$$C(e,s,0) = (W(e,s,0) - W^0) + \frac{(e+s)^2}{2(\delta-1)(x+1)}. \tag{36}$$

Because constraint (16c) implies $W(e,s,0) - W^0 \ge 0$, it follows from (36) that $C(e,s,0) \ge \frac{(e+s)^2}{2(\delta-1)(x+1)}$, which holds with equality only when constraint (16c) is binding. Further, because $x < \bar{x}$, $(d-p)-(r-d)x > 0$. It follows from (33) that for all (e, s),

$$[C(e,s,0)]^+ \geq C(e,s,0) \geq \frac{(e+s)^2}{2(\delta-1)(x+1)} \geq \frac{(\beta((d-p)-(r-d)x))^2}{2(\delta-1)(x+1)}, \tag{37}$$

Because $\frac{(\beta((d-p)-(r-d)x))^2}{2(\delta-1)(x+1)} \geq 0$, it follows that (37) holds with equality if and only if both constraints (16b) and (16c) are binding. □

Proof of Proposition 1. The proof proceeds in two steps. In Step 1, we find the optimal solution (e^*, s^*). In Step 2, we prove statements (i) and (ii).

Step 1. To solve for (e^*, s^*), we consider two cases based on the definition of $D_m(e,s,0)$ in (10), which implies $D_m(e,s,0) = \min\left\{1, \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}\right\}$: where (e,s) satisfies $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$, and where (e,s) satisfies $D_m(e,s,0) = 1$. Then, we show that the optimal cost in is lower in the first case.

Case 1: Suppose $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$. This condition is equivalent to

$$e + s \leq (\delta-1)(x+1) - ((d-p)-(r-d)x), \tag{38}$$

which we add to problem (16) as an additional constraint in this case. Because the objective is to minimize $[C(e,s,z)]^+$, to find an optimal solution it suffices to construct a feasible solution (e_1, s_1) that attains the lower bound in Lemma 1. Let (e_1, s_1) be given by the unique solution to the equations

$$D_m(e,s,0) - D_m^0 = \beta D_m^0,$$
$$W(e,s,0) = W^0,$$

which correspond to constraints (16b) and (16c). Solving for (e,s) yields the following unique solution:

$$e_1 = \frac{\beta(\beta+2)((d-p)-(r-d)x)^2}{2(\delta-1)(x+1)}$$
$$s_1 = \frac{\beta((d-p)-(r-d)x)(2(\delta-1)(x+1)-(\beta+2)((d-p)-(r-d)x))}{2(\delta-1)(x+1)}.$$

Next, we show that (e_1, s_1) is feasible. By construction, (e_1,s_1) satisfies constraints (16b) and (16c) with equality. Note that $e_1 + s_1 = \beta((d-p)-(r-d)x)$. Because $0 < \beta < \bar{\tau}(x)$, (e_1, s_1) satisfies constraint (38). Next, because $\delta > 1$ and $x > \underline{x}$, we have $e_1 > 0$. In addition, because $0 < \beta < \bar{\tau}(x)$, we also have $s_1 > 0$. Further, for any (e,s), the revenue *increment* of nearby merchants is given by $\Pi_s(e,s,0) - \Pi_s^0 = K\frac{e+s}{(\delta-1)(x+1)}$, which follows from (15). Because $e_1 > 0$ and $s_1 > 0$, it follows that $\Pi_s(e,s,0) - \Pi_s^0 > 0$, and thus (e_1, s_1) satisfies constraint (16d). Hence, (e_1, s_1) is feasible to (16). Because formulation (16) minimizes $[C(e,s,0)]^+$, and (e_1, s_1) achieves the lower bound of $[C(e,s,0)]^+$ presented in Lemma 1, it follows that (e_1, s_1) is optimal to (16) when restricting (e,s) to satisfy $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$.

Case 2: Suppose $D_m(e,s,0) = 1$. It follows that $D_d(e,s,0) = 0$ and $v(e,s,0) = 1$. Constraint (16c) can then be rewritten as

$$W(e,s,0) - W^0 = \int_0^1 U_m(e,s,0)dV - \left[\int_0^{v^0} U_m^0 dV + \int_{v^0}^1 U_d^0 dV\right] \tag{39}$$
$$= s - \frac{((\delta-1)(x+1) - ((d-p) - x(r-d)))^2}{2(\delta-1)(x+1)}. \tag{40}$$

Let $\phi = \frac{((\delta-1)(x+1)-((d-p)-x(r-d)))^2}{2(\delta-1)(x+1)}$. Then constraint (16c) requires that (e,s) satisfies $s \geq \phi$. Using (11) and the fact that $D_m(e,s,0) = 1$, we have $C(e,s,0) = s \geq 0$. Therefore, $[C(e,s,0)]^+ = C(e,s,0) \geq \phi$. Because $0 < \beta < \bar{\tau}(x)$, it is straightforward to verify that $\phi > C(e_1,s_1,0) = [C(e_1,s_1,0)]^+$. Hence, for all (e,s) that satisfies $D_m(e,s,0) = 1$, $[C(e,s,0)]^+ > [C(e_1,s_1,0)]^+$. Therefore, the optimal solution to (16), denoted by (e^*,s^*), must satisfy $D_m(e^*,s^*,0) = \frac{(d-p)-(r-d)x+e^*+s^*}{(\delta-1)(x+1)}$. It follows that $(e^*,s^*) = (e_1,s_1)$.

Step 2. Next, we prove statements (i) and (ii). (i). It is straightforward to verify that $\frac{\partial e^*}{\partial \beta} = \frac{(\beta+1)((d-p)-(r-d)x)^2}{(\delta-1)(x+1)} > 0$. Next, by taking the first derivative of e^* with respect to x, we have $\frac{\partial e^*}{\partial x} = \frac{\beta(\beta+2)\left((x+1)^2(r-d)^2-(r-p)^2\right)}{2(\delta-1)(x+1)^2}$. Because $x \in (\underline{x},\bar{x})$ and $r > d > p > 0$, it follows that $(x+1)^2(r-d)^2 < (r-d+d-p)^2 = (r-p)^2$, and thus $\frac{\partial e^*}{\partial x} < 0$. (ii). Note that $\frac{\partial s^*}{\partial \beta} = \frac{((d-p)-(r-d)x)((\delta-1)(x+1)-(\beta+1)((d-p)-(r-d)x))}{(\delta-1)(x+1)}$. Because $\delta > 1$, $\beta < \bar{\tau}(x)$ and $x \in (\underline{x},\bar{x})$, it follows that $\frac{\partial s^*}{\partial \beta} > 0$. To compute $\frac{\partial s^*}{\partial x}$, first note that $\frac{\partial^2 s^*}{\partial x^2} = -\frac{\beta(\beta+2)(r-p)^2}{(\delta-1)(x+1)^3} < 0$, meaning s^* is strictly concave in x and $\frac{\partial s^*}{\partial x}$ decreases in x. We next check the first derivatives at the boundaries, i.e. $\left.\frac{\partial s^*}{\partial x}\right|_{x=\bar{x}}$ and $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}}$. Note $\left.\frac{\partial s^*}{\partial x}\right|_{x=\bar{x}} = -\beta(r-d) < 0$. If $\delta < \delta_1 = 1+d-p$, then $\underline{x} = \frac{(d-p)-(\delta-1)}{(\delta-1)+(r-d)}$, and thus $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} = \frac{1}{2}\beta((\beta+2)(\delta-1)+2(\beta+1)(r-d)) > 0$; if $\delta \geq \delta_1$, then $\underline{x} = 0$, and thus $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} = -\frac{\beta(2(\delta-1)(r-d)-(\beta+2)(d-p)(2r-d-p))}{2(\delta-1)}$. Similarly, if $\delta < \delta_2 = \frac{(\beta+2)(d-p)(2r-d-p)}{2(r-d)} + 1$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=0} > 0$; if $\delta \geq \delta_2$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=0} \leq 0$. Because $\delta_2 - \delta_1 = \frac{\beta(d-p)(2r-d-p)+2(d-p)(r-p)}{2(r-d)} > 0$, we have $\delta_2 > \delta_1$. Therefore, if $\delta \geq \delta_2$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} \leq 0$, and if $1 < \delta < \delta_2$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} > 0$. Therefore, by continuity of $\frac{\partial s^*}{\partial x}$, if $\delta \geq \delta_2$, then s^* strictly decreases in $x \in (\underline{x},\bar{x})$. If $\delta < \delta_2$, then there exists $\tilde{x} \in (\underline{x},\bar{x})$ such that s^* strictly increases in x on $(\underline{x},\tilde{x})$ and strictly decreases in x on $[\tilde{x},\bar{x})$. □

Proof of Corollary 1. Using the expressions for (e^*,s^*) in Proposition 1, we have

$$e^* - s^* = \beta((d-p)-(r-d)x)\left(\frac{(\beta+2)((d-p)-(r-d)x)}{(\delta-1)(x+1)} - 1\right).$$

Because $x \in (\underline{x},\bar{x})$, $(d-p)-(r-d)x > 0$. Further, because $\delta > 1$, we have $e^* \geq s^*$ if and only if $x \leq \bar{x} - \frac{(\delta-1)(r-p)}{(r-d)((\beta+2)(r-d)+\delta-1)}$. The result follows by setting $\hat{x} = \bar{x} - \frac{(\delta-1)(r-p)}{(r-d)((\beta+2)(r-d)+\delta-1)}$. Next, note that because $\delta > 1$, $r > p$ and $r > d$, it follows that $\hat{x} < \bar{x}$. Define $\bar{\delta} = (\beta+2)(d-p)+1$, which satisfies $\bar{\delta} > 1$. It remains to show that $\underline{x} \leq \hat{x}$ if and only if $\delta \geq \bar{\delta}$. From the definition of $\underline{x}$, if $\delta < 1+d-p$, then $\underline{x} = \frac{(d-p)-(\delta-1)}{(\delta-1)+(r-d)}$. It is straightforward to check that $\hat{x} \in (\underline{x},\bar{x})$. If $\delta \geq 1+d-p$, then $\underline{x} = 0$, and thus $\hat{x} \leq \underline{x}$ only if $\delta \geq \bar{\delta}$, where $\bar{\delta} > 1+d-p$. Therefore, $s^* > e^*$ holds for all $x \in (\underline{x},\bar{x})$ if and only if $\hat{x} \leq \underline{x}$, which holds if and only if $\delta \geq \bar{\delta}$.

Proof of Corollary 2. Using the expressions for e^* and s^* from Proposition 1 and the cost function given in (11), the total cost under (e^*,s^*) is $C(e^*,s^*,0) = \frac{\beta^2((d-p)-(r-d)x)^2}{2(\delta-1)(x+1)}$. Because $x < \bar{x}$ and $\delta > 1$,

$C(e^*, s^*, 0) > 0$. By noting that $\frac{\partial C(e^*,s^*,0)}{\partial \beta} = \frac{\beta((d-p)-x(r-d))^2}{(\delta-1)(x+1)} > 0$, we conclude that $C(e^*, s^*, 0)$ increases in β. Next, note $\frac{\partial C(e^*,s^*,0)}{\partial x} = \frac{\beta^2\left((x+1)^2(r-d)^2-(r-p)^2\right)}{2(\delta-1)(x+1)^2}$. Because $x \in (\underline{x}, \bar{x})$, we have $0 < (x+1)(r-d) < (r-p)$. It follows that $\frac{\partial C(e^*,s^*,0)}{\partial x} < 0$. □

Proof of Proposition 2. From the definition in (11), we have $[C(0,0,z)]^+ = C(0,0,z) = 0$. Thus, under mechanism [I], the operating cost is always 0. It follows that any feasible solution to problem (18) is optimal. Therefore, solving for the optimal subsidy z^* is equivalent to finding the optimal solution to the following problem:

$$\max_{z \geq 0} \Pi_s(0,0,z) \tag{41a}$$

$$\text{s.t. } D_m(0,0,z) - D_m^0 \geq \beta D_m^0 \tag{41b}$$

$$W(0,0,z) \geq W^0 \tag{41c}$$

$$\Pi_s(0,0,z) \geq \Pi_s^0 \tag{41d}$$

The proof is composed of three steps. In Step 1, we prove (41) is feasible if and only if $\beta \leq \tau(x)$, which is also a sufficient and necessary condition for formulation (18) to be feasible. Next, in Step 2, we solve for the optimal solution z^* to (41), which is also optimal to (18). Finally, in Step 3, we conduct the comparative statics analysis of z^*.

Step 1. To construct the sufficient and necessary condition for problem (41) to be feasible, we establish the conditions under which the feasible solution set is non-empty. Let us first define the feasible set. Let $\tilde{z} = (\delta-1)(x+1) - ((d-p)-(r-d)x)$. To find all the feasible solutions to problem (41), we consider two cases based on the definition of $D_m(0,0,z)$ in (10), which implies $D_m(0,0,z) = \min\left\{1, \frac{(d-p)-(r-d)x+z}{(\delta-1)(x+1)}\right\}$: $D_m(0,0,z) = \frac{(d-p)-(r-d)x+z}{(\delta-1)(x+1)}$ and $D_m(0,0,z) = 1$. In preparation, note that because $\underline{x} < x < \bar{x}$ and $\delta > 1$, we have $(\delta-1)(x+1) > ((d-p)-x(r-d)) > 0$.

<u>Case 1:</u> Suppose $D_m(0,0,z) = \frac{(d-p)-(r-d)x+z}{(\delta-1)(x+1)}$, which is equivalent to $z \leq \tilde{z}$. In this case, we first show that any $z \geq 0$ satisfies constraint (41c). Using the welfare expression given in (12), the commuter welfare increment is $W(0,0,z) - W^0 = \frac{z(2(d-p)-2(r-d)x+z)}{2(\delta-1)(x+1)}$. Because $\delta > 1$ and $x \in (\underline{x}, \bar{x})$, we have $W(0,0,z) - W^0 \geq 0$ if $z \geq 0$. Therefore, constraint (41c) is satisfied for all $z \geq 0$. Next, we show that when restricting $z \leq \tilde{z}$, the feasible region of problem (41) is $[\underline{z}, \bar{z}]$, where $\underline{z} = \beta((d-p)-x(r-d))$ and $\bar{z} = \min\{K, (\delta-1)(x+1)\} - ((d-p)-x(r-d))$. To do so, we prove that constraint (41b) is satisfied if and only if $z \geq \underline{z}$, and constraint (41d) is satisfied if and only if $0 \leq z \leq K - ((d-p)-x(r-d))$. First, using the definition of D_m^0 in (3), it is straightforward to verify that constraint (41b) holds if and only if $\frac{z}{(\delta-1)(x+1)} \geq \beta\frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}$. Because $\delta > 1$, it follows that constraint (41b) holds if and only if $z \geq \underline{z} = \beta((d-p)-x(r-d))$. Because $x \leq \bar{x}$, it follows that $\underline{z} > 0$. Next, using the revenue expression in (15), when restricting $z \leq \tilde{z}$, it is straightforward to verify that constraint (41d) holds if and only if $0 \leq z \leq K - ((d-p)-x(r-d))$. Therefore, when adding the additional constraint $z \leq \tilde{z}$ to problem (41), the feasible region is $[\underline{z}, \min\{\tilde{z}, K - ((d-p)-x(r-d))\}]$, or equivalently, $[\underline{z}, \bar{z}]$.

<u>Case 2:</u> Suppose $D_m(0,0,z) = 1$, which is equivalent to $z \geq \tilde{z}$. Let $\dot{z} = K(1 - \frac{(d-p)-(r-d)x}{(\delta-1)(x+1)})$. It is straightforward to verify that constraint (41d) is satisfied if and only if $z \leq \dot{z}$. Next, we show that

constraints (41b) and (41c) are satisfied if $z \geq \tilde{z}$. First, note that

$$D_m(0,0,z) - D_m^0 = 1 - \frac{(d-p)-(r-d)x}{(\delta-1)(x+1)} \geq \beta\frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}$$

holds for $z \geq \tilde{z}$ because $\beta < \bar{\tau}(x)$; thus constraint (41b) is satisfied for all $z \geq \tilde{z}$. Next, from the definition in (10), if $z \geq \tilde{z}$, then

$$\begin{aligned}W(0,0,z) - W^0 &= \frac{(\delta-1)(x+1)((x+1)(2d-\delta)-2p-2rx+2z+x+1)-(-d(x+1)+p+rx)^2}{2((\delta-1)(x+1))} \\ &> \frac{(\delta-1)^2(x+1)^2-((d-p)-x(r-d))^2}{2((\delta-1)(x+1))} \\ &\geq 0,\end{aligned}$$

which implies constraint (41c) is satisfied for all $z \geq \tilde{z}$. Therefore, constraint (41d) is satisfied if and only if $z \leq \dot{z}$, and constraints (41b) and (41c) are satisfied if $z \geq \tilde{z}$.

To summarize Case 1 and Case 2, note that $\bar{z} \leq \tilde{z}$, and thus the feasible region of problem (41) is $[\underline{z},\bar{z}] \cup [\tilde{z},\dot{z}]$. Now that we have constructed the feasible set, next we prove $[\underline{z},\bar{z}] \cup [\tilde{z},\dot{z}]$ is non-empty if and only if $\beta \leq \tau(x)$. The derivation consists of two cases: $K < (\delta-1)(x+1)$ and $K \geq (\delta-1)(x+1)$. If $K < (\delta-1)(x+1)$, it is straightforward to verify that $\dot{z} < \tilde{z}$, in which case the feasible region is $[\underline{z},\bar{z}]$; note that this interval is non-empty if and only if $\underline{z} \leq \bar{z}$, or equivalently, $\beta \leq \frac{K}{((d-p)-(r-d)x)} - 1$. If $K \geq (\delta-1)(x+1)$, then $\bar{z} = \tilde{z} \leq \dot{z}$, in which case the feasible region is $[\underline{z},\dot{z}]$; note that this interval is non-empty if and only if $\underline{z} \leq \dot{z}$, or equivalently, $\beta \leq \frac{K}{(\delta-1)(x+1)}\bar{\tau}(x)$. Further, in this case where $K \geq (\delta-1)(x+1)$, it holds that $\bar{\tau}(x) \leq \frac{K}{(\delta-1)(x+1)}\bar{\tau}(x)$. To summarize, because $\beta < \bar{\tau}(x)$, the feasible region $[\underline{z},\bar{z}] \cup [\tilde{z},\dot{z}]$ is non-empty if and only if $\beta \leq \min\{\frac{K}{((d-p)-(r-d)x)} - 1, \bar{\tau}(x)\}$. By definition, $\tau(x) = \min\{\frac{K}{((d-p)-(r-d)x)} - 1, \bar{\tau}(x)\}$. Because $K > \underline{K}$, it follows that $\tau(x) > 0$. Therefore, (41) is feasible if and only if $\beta \leq \tau(x)$, which is also sufficient and necessary for formulation (18) to be feasible.

Step 2. In this step, we solve for the optimal solution z^*. To do so, we first show that problem (41) is equivalent to the following problem:

$$\max \Pi_s(0,0,z) \text{ s.t. } z \in [\underline{z},\bar{z}]. \tag{42}$$

To do so, we consider two cases: $K < (\delta-1)(x+1)$ and $K \geq (\delta-1)(x+1)$. If $K < (\delta-1)(x+1)$, we have showed in Step 1 that the optimal solution z^* lies in the feasible region and satisfies $z^* \in [\underline{z}, K-((d-p)-x(r-d))]$. If $K \geq (\delta-1)(x+1)$, it follows from Step 1 that for any β that satisfies $0 < \beta \leq \tau(x)$, the feasible region is $[\underline{z},\dot{z}]$, where $\dot{z} = K\left(1 - \frac{(d-p)-(r-d)x}{(\delta-1)(x+1)}\right) \geq \tilde{z}$. Note that when $z \in [\tilde{z},\dot{z}]$, from the definition in (15), we have $\Pi_s(0,0,z) = K - z$, which decreases in z. It follows that $\Pi_s(0,0,\tilde{z}) > \Pi_s(0,0,z)$ for all $z > \tilde{z}$. Therefore, when $K \geq (\delta-1)(x+1)$, to maximize $\Pi_s(0,0,z)$, the optimal solution z^* should satisfy $z^* \in [\underline{z},\tilde{z}]$, where $\tilde{z} = (\delta-1)(x+1) - ((d-p)-(r-d)x)$. To conclude, z^* must satisfy $z^* \in [\underline{z}, \min\{K,(\delta-1)(x+1)\} - ((d-p)-x(r-d))]$, or equivalently, $z^* \in [\underline{z},\bar{z}]$. Thus formulation (41) is equivalent to formulation (42). Next, we solve problem (42). Using the profit expression given in (15), and the fact that $\delta > 1$, we have $\frac{\partial^2\Pi_s(0,0,z)}{\partial z^2} = -\frac{2}{(\delta-1)(x+1)} < 0$. Hence, $\Pi_s(0,0,z)$ is strictly concave in z. The solution to the first order condition $\frac{\partial \Pi_s(0,0,z)}{\partial z} = 0$ is $\hat{z} = \frac{K}{2} - \frac{1}{2}((d-p)-(r-d)x)$. Therefore, the unique optimal subsidy z^* is given by

$$z^* = \begin{cases} \underline{z} & \text{if } \hat{z} < \underline{z}, \\ \hat{z} & \text{if } \underline{z} \le \hat{z} \le \bar{z}, \\ \bar{z} & \text{if } \bar{z} < \hat{z}, \end{cases}$$

or, equivalently, $z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\}$. Note that z^* is the unique optimal solution to problem (42), so z^* is the unique feasible, and thus optimal, solution to (18).

Step 3. Because $\bar{z} = \min\{K, (\delta-1)(x+1)\} - ((d-p) - x(r-d))$, we can rewrite z^* as $z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}_1, \bar{z}_2\}$, where $\bar{z}_1 = K - ((d-p) - x(r-d))$ and $\bar{z}_2 = (\delta-1)(x+1) - ((d-p) - x(r-d))$. Note $\frac{\partial \underline{z}}{\partial x} = -\beta(r-d) < 0$, $\frac{\partial \hat{z}}{\partial x} = \frac{r-d}{2} > 0$, $\frac{\partial \bar{z}_1}{\partial x} = r-d > 0$ and $\frac{\partial \bar{z}_2}{\partial x} = (\delta-1)+(r-d) > 0$. Because $\underline{z}, \hat{z}, \bar{z}_1$, and $\bar{z}_2$ are all continuous in x, z^* is continuous in x. Therefore, if $z^* = \underline{z}$, z^* decreases in x; otherwise, z^* increases in x. It remains to show the conditions under which $z^* = \underline{z}$. In preparation, let $\tilde{x} = \bar{x} - \frac{K}{(2\beta+1)(r-d)}$ and $K_1 = (\bar{x} - \underline{x})(2\beta+1)(r-d)$, and note that $\tilde{x} < \bar{x}$ by definition of $\tilde{x}$ and because $r > d$. Note that $z^* = \underline{z}$ if and only if $\hat{z} \le \underline{z}$, which holds if and only if $x \le \tilde{x}$. Because $x \in (\underline{x}, \bar{x})$, $x \le \tilde{x}$ can hold only if $\tilde{x} > \underline{x}$. It can be verified that $\tilde{x} > \underline{x}$ if and only if $K < K_1$. Note that $K_1 = (\bar{x} - \underline{x})(2\beta+1)(r-d) \ge (d-p) - (r-d)\underline{x} > (d-p) - (r-d)x = \underline{K}$. Therefore, (i) if $K \ge K_1$, z^* increases in x on $(\underline{x}, \bar{x})$; (ii) if $\underline{K} < K < K_1$, $z^* = \underline{z}$ decreases in x on $(\underline{x}, \tilde{x}]$ and increases in x on $(\tilde{x}, \bar{x})$. The result follows. □

Proof of Theorem 1. First, we compare the performances of the two mechanisms in metrics I – V. By Corollary 2, we have $C^D = C(e^*, s^*, 0) > 0$, and under mechanism [I], $C^I = C(0, 0, z^*) = 0$. Hence, $C^I = 0 < C^D$. Next, using the expressions for e^* and s^* given in Proposition 1, and the expression for z^* given in Proposition 2, it can be shown that $D_m(0,0,z^*) - D_m^0 \ge \beta D_m^0 = D_m(e^*, s^*, 0) - D_m^0$, or equivalently, $D_m^I \ge D_m^D$. Because $D_m^I \ge D_m^D$, from (13) and (14) we immediately obtain $\Pi_p^I \ge \Pi_p^D$ and $\Pi_r^I \ge \Pi_r^D$, respectively. To compare the commuter welfare in mechanism [D] and mechanism [I], it is easy to check that $W^D = W^0$, because constraint (16c) is binding in equilibrium. Further, $W^I - W^0 = W(0,0,z^*) - W^0 = \frac{z^*(2(d-p) - 2(r-d)x + z^*)}{2(\delta-1)(x+1)}$, which satisfies $W^I - W^0 > 0$ if and only if $z^* > 0$. To see that $z^* > 0$ holds, let us first define $\gamma_1 = (\delta-1)(x+1) > 0$ and $\gamma_2 = (d-p) - x(r-d) > 0$. Then $\underline{z} = \beta\gamma_2 > 0$, $\hat{z} = \frac{K-\gamma_2}{2}$, $\bar{z} = \min\{K - \gamma_2, \gamma_1 - \gamma_2\}$ and $\tau(x) = \min\{\frac{K}{\gamma_2} - 1, \frac{\gamma_1}{\gamma_2} - 1\}$. Because $K > \underline{K} = \gamma_2$, it follows $\hat{z} = \frac{K-\gamma_2}{2} > 0$ and $\bar{z} = \min\{K - \gamma_2, \gamma_1 - \gamma_2\} > 0$. Therefore, $z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\} > 0$, and thus $W^I > W^0 = W^D$.

For the enterprise's profit, using the definition in (15), we have $\Pi_s^I - \Pi_s^D = \frac{\hat{z}^2 + K\gamma_2 - (z^* - \hat{z})^2}{\gamma_1} - \frac{K\gamma_2(\beta+1)}{\gamma_1}$. Define $\tilde{\tau}(x) = \frac{\hat{z}^2 - (\hat{z} - z^*)^2}{K\gamma_2}$. Then $\Pi_s^I < \Pi_s^D$ holds if and only if $\beta > \tilde{\tau}(x)$. Note mechanism [I] is feasible if and only if $\beta \in (0, \tau(x)]$, by Proposition 2. To see there exists β such that both $\beta > \tilde{\tau}(x)$ holds and mechanism [I] is feasible, it suffices to show $0 \le \tilde{\tau}(x) \le \tau(x)$ for all x. Because $z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\} \in \{\underline{z}, \hat{z}, \bar{z}\}$, we shall prove $0 \le \tilde{\tau}(x) \le \tau(x)$ for all x by considering three cases: $z^* = \underline{z}$, $z^* = \hat{z}$ and $z^* = \bar{z}$. Because $0 < \beta \le \tau(x)$, it follows that $\underline{z} \le \bar{z}$, and thus the three cases correspond to $\hat{z} \le \underline{z}$, $\underline{z} < \hat{z} \le \bar{z}$, and $\bar{z} < \hat{z}$.

Case 1: Suppose $\hat{z} \le \underline{z}$. Then $z^* = \underline{z}$. It follows that $\tilde{\tau}(x) = \frac{\hat{z}^2 - (\hat{z} - \underline{z})^2}{K\gamma_2} = \frac{2\underline{z}\hat{z} - \underline{z}^2}{K\gamma_2} \le \frac{\underline{z}^2}{K\gamma_2} \le \frac{\bar{z}^2}{K\gamma_2} = \frac{(\min\{K,\gamma_1\} - \gamma_2)^2}{K\gamma_2} < \frac{(\min\{K,\gamma_1\} - \gamma_2)\min\{K,\gamma_1\}}{K\gamma_2} = \tau(x)\min\{1, \frac{\gamma_1}{K}\} \le \tau(x)$. Further, because $\beta \le \tau(x) = \min\left\{\frac{K}{\gamma_2} - 1, \frac{\gamma_1}{\gamma_2} - 1\right\}$, it follows that $\beta\gamma_2 \le K - \gamma_2$, or equivalently, $z^* = \underline{z} \le 2\hat{z}$. Thus $\tilde{\tau}(x) = \frac{\hat{z}^2 - (\underline{z} - \hat{z})^2}{\gamma_1\gamma_2} \ge 0$. Therefore, in this case, $0 \le \tilde{\tau}(x) \le \tau(x)$ holds for all x.

Case 2: Suppose $\underline{z} < \hat{z} \le \bar{z}$. Then $z^* = \hat{z}$ and $\tilde{\tau}(x) = \frac{\hat{z}^2}{K\gamma_2} \le \frac{\bar{z}^2}{K\gamma_2}$. In Case 1 we show that $\tilde{\tau}(x) \le \frac{\bar{z}^2}{K\gamma_2} <$

$\tau(x)$. Further, it is straightforward to verify that $\tilde{\tau}(x) > 0$. Therefore, $0 \le \tilde{\tau}(x) \le \tau(x)$ holds for all x.

Case 3: Suppose $\bar{z} < \hat{z}$. Then $z^* = \bar{z}$. Because $\bar{z} = \min\{\gamma_1 - \gamma_2, K - \gamma_2\}$ and $\hat{z} = \frac{K-\gamma_2}{2} < K - \gamma_2$, $\bar{z} \le \hat{z}$ implies $\bar{z} = \gamma_1 - \gamma_2 < K - \gamma_2$, or equivalently, $\gamma_1 < K$. It follows that $\tau(x) = \frac{\gamma_1}{\gamma_2} - 1$, and thus $\tilde{\tau}(x) = \frac{2\bar{z}\hat{z} - \bar{z}^2}{K\gamma_2} = \frac{2(\gamma_1-\gamma_2)\frac{K-\gamma_2}{2} - (\gamma_1-\gamma_2)^2}{K\gamma_2} = \tau(x)(1 - \frac{\gamma_1}{K}) < \tau(x)$. It is straightforward to verify that $\tilde{\tau}(x) > 0$. Therefore, $0 \le \tilde{\tau}(x) \le \tau(x)$ holds for all x.

To summarize the three cases above, (i) if $\beta \le \tilde{\tau}(x)$, then $\Pi_s^I \ge \Pi_s^D$; (ii) if $\beta \in (\tilde{\tau}(x), \tau(x)]$, then $\Pi_s^I < \Pi_s^D$. The result follows.

Proof of Proposition 3. Similar to the proof of Proposition 1, the proof proceeds in two steps. In Step 1, we find the optimal solutions (e^*, s^*). In Step 2, we prove the comparative statics results.

Step 1. To solve for (e^*, s^*) , we consider two cases: $(D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)})$ and $(D_m(e,s,0) = 1)$. We then show that the optimal commuter welfare is higher in the latter case.

Case 1: Suppose $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$, which is equivalent to

$$e + s \le (\delta-1)(x+1) - ((d-p) - (r-d)x). \tag{43}$$

The proof proceeds in two steps. First, we construct an upper bound for the optimal value of the objective function (20a) when adding (43) as an additional constraint to problem (20); second, to show optimality, we construct a feasible solution that attains this bound. First, because $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$, we can apply (36) and rewrite constraint (20c) as

$$C(e,s,0) = (W(e,s,0) - W^0) + \frac{(e+s)^2}{2(\delta-1)(x+1)} = 0,$$

which implies any feasible (e,s) should satisfy $W(e,s,0) - W^0 = -\frac{(e+s)^2}{2(\delta-1)(x+1)}$. Next, constraint (20b) can be rewritten as $e + s \ge \beta((d-p) - (r-d)x)$. Note $\beta((d-p) - (r-d)x) > 0$ because $x \in (\underline{x}, \bar{x})$. Therefore,

$$W(e,s,0) - W^0 = -\frac{(e+s)^2}{2(\delta-1)(x+1)} \le -\frac{(\beta((d-p) - x(r-d)))^2}{2(\delta-1)(x+1)},$$

which holds with equality if and only if constraint (20b) is binding. Next, we prove that the upper bound is attainable. We first solve for (e,s) such that constraint (20b) is binding and constraint (20c) is satisfied:

$$\begin{aligned} &D_m(e,s,0) - D_m^0 = \beta D_m^0 \\ &C(e,s,0) = 0. \end{aligned} \tag{44}$$

Solving the above set of equations yields the unique solution

$$\begin{aligned} e_2 &= \frac{\beta(\beta+1)((d-p) - (r-d)x)^2}{(\delta-1)(x+1)} \\ s_2 &= \frac{\beta((d-p) - (r-d)x)((\delta-1)(x+1) - (\beta+1)((d-p) - (r-d)x))}{(\delta-1)(x+1)}. \end{aligned}$$

Using similar arguments in proof of Proposition 1, it is plain to verify that (e_2, s_2) satisfies the additional constraint (43) and it's feasible to problem (20), so (e_2, s_2) is the unique optimal solution to formulation (20) when restricting $D_m(e,s,0) = \frac{(d-p)-(r-d)x+e+s}{(\delta-1)(x+1)}$. The corresponding optimal commuter welfare increment is $\Delta W(e_2, s_2, 0) = -\frac{(\beta((d-p)-x(r-d)))^2}{2(\delta-1)(x+1)}$.

Case 2: Suppose $D_m(e,s,0) = 1$, which is equivalent to $e+s \geq (\delta-1)(x+1) - ((d-p)-(r-d)x)$. It follows that $D_d(e,s,0) = 0$. Then the budget neutrality constraint (20c) can be rewritten as

$$C(e,s,0) = sD_m(e,s,0) - eD_d(e,s,0) = s = 0. \tag{45}$$

From $D_m(e,s,0) = 1$, it follows that $e \geq (\delta-1)(x+1) - ((d-p)-(r-d)x)$. From the definition in (12), the commuter welfare increment is

$$\Delta W(e,s,0) = -\frac{((\delta-1)(x+1) - ((d-p)-x(r-d)))^2}{2(\delta-1)(x+1)}. \tag{46}$$

Because $0 < \beta < \bar{\tau}(x)$, it can be verified that for all (e,s) that satisfies $D_m(e,s,0) = 1$, $\Delta W(e,s,0) < \Delta W(e_2, s_2, 0)$. Therefore, the unique optimal solution to (20), denoted by (e^*, s^*), must satisfy $D_m(e^*, s^*, 0) = \frac{(d-p)-(r-d)x+e^*+s^*}{(\delta-1)(x+1)}$ and $(e^*, s^*) = (e_2, s_2)$.

Step 2. Next, we prove statements (i)-(iii) in order. (i). We first take the first derivative of e^* with respect to x and β. Because $\delta > 1$, $0 < x < \bar{x}$ and $r > d > p > 0$, it follows that $\frac{\partial e^*}{\partial \beta} = \frac{(2\beta+1)((d-p)-(r-d)x)^2}{(\delta-1)(x+1)} > 0$. Next, $\frac{\partial e^*}{\partial x} = \frac{\beta(\beta+1)\left((x+1)^2(r-d)^2-(r-p)^2\right)}{(\delta-1)(x+1)^2}$. Because $r-p = (d-p)+(r-d) > (r-d)(x+1)$, it follows that $\frac{\partial e^*}{\partial x} < 0$. Therefore, e^* strictly increases in in β and strictly decreases in x. (ii). First, note that s^* is concave in x: $\frac{\partial^2 s^*}{\partial x^2} = -\frac{2\beta(\beta+1)(p-r)^2}{(\delta-1)(x+1)^3} < 0$. Then, to show the result, it suffices to check the first derivatives at the boundaries: $\left.\frac{\partial s^*}{\partial x}\right|_{x=\bar{x}}$ and $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}}$. If $\delta < \delta_1 = 1+d-p$, then $\underline{x} = \frac{(d-p)-(\delta-1)}{(\delta-1)+(r-d)}$, and $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} = \beta((\beta+1)(\delta-1)+(2\beta+1)(r-d)) > 0$; if $\delta \geq \delta_1$, then $\underline{x} = 0$, and $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} = \frac{\beta((\beta+1)(d-p)(2r-d-p)-(\delta-1)(r-d))}{\delta-1}$. If $\delta < \delta_3 = \frac{(\beta+1)(d-p)(2r-d-p)}{r-d} + 1$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} > 0$; if $\delta \geq \delta_3$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} \leq 0$. Notice $\delta_3 - \delta_1 = \frac{(d-p)(\beta(2r-d-p)+r-p)}{r-d} > 0$, or $\delta_3 > \delta_1$. We conclude that if $\delta \geq \delta_3$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} \leq 0$, and if $1 < \delta < \delta_3$, $\left.\frac{\partial s^*}{\partial x}\right|_{x=\underline{x}} > 0$. Therefore, by continuity of $\frac{\partial s^*}{\partial x}$ in x, it follows that if $\delta \geq \delta_3$, s^* decreases in $x \in (\underline{x}, \bar{x})$, and if $\delta < \delta_3$, there exists $\tilde{x} \in (\underline{x}, \bar{x})$ such that s^* strictly increases in x on $(x, \tilde{x})$ and strictly decreases in x on $[\tilde{x}, \bar{x})$. (iii). Note $\frac{\partial s^*}{\partial \beta} = \frac{1}{(\delta-1)(x+1)}(((d-p)-(r-d)x)((\delta-1)(x+1)-(2\beta+1)((d-p)-(r-d)x)))$. Then if $\beta < \frac{(\delta-1)(x+1)}{2((d-p)-(r-d)x)} - \frac{1}{2} = \frac{\bar{\tau}(x)}{2}$, $\frac{\partial s^*}{\partial \beta} > 0$, and if $\beta \geq \frac{\bar{\tau}(x)}{2}$, $\frac{\partial s^*}{\partial \beta} \leq 0$. Therefore, s^* increases in $\beta \in (0, \frac{\bar{\tau}(x)}{2})$ and decreases in $\beta \in [\frac{\bar{\tau}(x)}{2}, \bar{\tau}(x))$. □

Proof of Corollary 3. Note $e^* - s^* = \beta((d-p)-(r-d)x)\left(\frac{2(\beta+1)((d-p)-(r-d)x)}{(\delta-1)(x+1)} - 1\right)$. Because $x \in (\underline{x}, \bar{x})$, $(d-p)-(r-d)x > 0$. Therefore, $e^* \geq s^*$ if and only if $\frac{2(\beta+1)((d-p)-(r-d)x)}{(\delta-1)(x+1)} - 1 \geq 0$. It follows that $e^* \geq s^*$ if and only if $x \leq \bar{x} - \frac{(\delta-1)(r-p)}{(r-d)(2(\beta+1)(r-d)+\delta-1)}$. Defining $\hat{x} = \bar{x} - \frac{(\delta-1)(r-p)}{(r-d)(2(\beta+1)(r-d)+\delta-1)}$ yields the main result. Following a parallel argument to the proof of Corollary 1, define $\bar{\bar{\delta}} = 2(\beta+1)(d-p)+1$, which satisfies $\bar{\bar{\delta}} > 1$. It can then be shown that $s^* > e^*$ for all $x \in (\underline{x}, \bar{x})$ if and only if $\hat{x} \leq \underline{x}$, which holds if and only if $\delta \geq \bar{\bar{\delta}}$. □

Proof of Corollary 4. By plugging in the optimal solutions (e^*, s^*) in Proposition 3, the change in commuter welfare is given by

$$\Delta W(e^*, s^*, 0) = W(e^*, s^*, 0) - W^0 = -\frac{\beta^2((d-p)-(r-d)x)^2}{2(\delta-1)(x+1)} < 0.$$

Note $\frac{\partial \Delta W(e^*,s^*,0)}{\partial \beta} = -\frac{\beta((d-p)-x(r-d))^2}{(\delta-1)(x+1)} < 0$, and thus $\Delta W(e^*, s^*, 0)$ strictly decreases in β. Next, note $\frac{\partial \Delta W(e^*,s^*,0)}{\partial x} = \frac{\beta^2\left((r-p)^2-(x+1)^2(r-d)^2\right)}{2(\delta-1)(x+1)^2}$. Because $x \in (\underline{x}, \bar{x})$, $0 < (x+1)(r-d) < (r-p)$. It follows that $\frac{\partial \Delta W(e^*,s^*,0)}{\partial x} > 0$, and thus $\Delta W(e^*, s^*, 0)$ strictly increases in x. $\square$

Proof of Proposition 4. Based on the proof of Proposition 2, constraints (18b), (18d) and (18e), which are equivalent to constraints (21b), (21d) and (21e), uniquely define a feasible and thus optimal solution $z^* = \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\}$. Because $C(0, 0, z^*) = 0$, it follows that z^* is also the unique feasible and thus optimal solution to formulation (21). To show that $W(0, 0, z) > W^0$, note that because $z^* > 0$, using the definition in (12), we have $W(0, 0, z) - W^0 = \frac{z(2(d-p)-2(r-d)x+z)}{2(\delta-1)(x+1)} > 0$. $\square$

Proof of Theorem 2. First, we compare the performance of mechanisms [D-A] and [I-A] in metrics I – V. By constraints (20c) and (21c), we have $C^I = C^D = 0$. Next, using the expressions for e^* and s^* given in Proposition 3, and the expression for z^* given in Proposition 4, it can be shown that $D_m(0, 0, z^*) - D_m^0 \geq \beta D_m^0 = D_m(e^*, s^*, 0) - D_m^0$, or equivalently, $D_m^I \geq D_m^D$. Because $D_m^I \geq D_m^D$, from (13) and (14) we immediately obtain $\Pi_p^I \geq \Pi_p^D$ and $\Pi_r^I \geq \Pi_r^D$, respectively. Recall from Corollary 4 and Proposition 4 that, $W(e^*, s^*, 0) - W^0 < 0$ and $W(0, 0, z^*) - W^0 > 0$, and thus $W^D < W^0 < W^I$. For metric VI, using definition in (15), we have $\Pi_s^I - \Pi_s^D = \frac{\hat{z}^2+K((d-p)-x(r-d))-(z^*-\tilde{z})^2}{(\delta-1)(x+1)} - \frac{K((d-p)-x(r-d))(\beta+1)}{(\delta-1)(x+1)}$, which is equal to the enterprise's profit difference between mechanisms [D] and [I]. Following the proof of Theorem 1, if $\beta \leq \tilde{\tau}(x)$, then $\Pi_s^I \geq \Pi_s^D$, and if $\beta \in (\tilde{\tau}(x), \tau(x)]$, then $\Pi_s^I < \Pi_s^D$. The result follows. $\square$

Before proving Proposition 5, we introduce two supporting results: Lemma 2 and Lemma 3. For ease of exposition, let $\gamma_1 = (\delta-1)(x+1)$ and $\gamma_2 = (d-p) - x(r-d)$, and note $\gamma_1 > 0$ and $\gamma_2 > 0$.

Lemma 2. *Suppose the constraint $D_m(e, s, z) = \frac{\gamma_2+e+s+z}{\gamma_1}$, or equivalently, $e + s + z \leq \gamma_1 - \gamma_2$, is added to formulation (22). Then formulation (22) is equivalent to*

$$\min_{e,s,z\geq 0} \ [C(e, s, z)]^+ \tag{47a}$$

$$s.t.\ z = \min\{\max\{z_1(e, s), z_2(e, s), z_3(e, s), 0\}, z_4(e, s), z_5(e, s)\} \tag{47b}$$

$$e \leq \frac{\gamma_1^2 - \gamma_2^2}{2\gamma_1} \tag{47c}$$

$$e + s \geq \frac{{\gamma_2}^2 + 2\gamma_1 e + \gamma_2 K}{\sqrt{\gamma_2^2 + 2\gamma_1 e}} - (\gamma_2 + K) \tag{47d}$$

$$e + s \leq \gamma_1 - \gamma_2 \tag{47e}$$

$$e + s \geq \frac{\beta((\beta+1)\gamma_2 - K)}{\beta+1}, \tag{47f}$$

where $z_1(e, s) = \frac{1}{2}(K - (\gamma_2 + (e + s)))$, $z_2(e, s) = \beta\gamma_2 - (e + s)$, $z_3(e, s) = \sqrt{\gamma_2^2 + 2\gamma_1 e} - (\gamma_2 + e + s)$, $z_4(e, s) = \gamma_1 - \gamma_2 - (e + s)$, and $z_5(e, s) = \sqrt{\left(\frac{K}{2} - \frac{1}{2}(\gamma_2 + (e + s))\right)^2 + K(e + s)} + \frac{K}{2} - \frac{1}{2}(\gamma_2 + (e + s))$.

Proof. Note that formulations (22) and (47) have the same objective functions. Therefore, to establish the equivalence between the two formulations, it suffices to prove that the feasible regions are equivalent. First, we show that when adding an additional constraint $e+s+z \leq \gamma_1 - \gamma_2$, which is equivalent to $z \leq z_4(e,s)$, the feasible region of problem (22) is $\mathcal{P}_1 = \{(e,s,z) |\max\{z_2(e,s), z_3(e,s), 0\} \leq z \leq \min\{z_4(e,s), z_5(e,s)\}, z = \min\{\max\{z_1(e,s), z_2(e,s), \ \ z_3(e,s), 0\}, z_4(e,s), z_5(e,s)\}, e, s, z \geq 0\}$. Because we assume (e,s,z) satisfies $D_m(e,s,z) = \frac{\gamma_2+e+s+z}{\gamma_1}$, it follows that $D_m(e,s,z) - D_m^0 = \frac{e+s+z}{\gamma_1}$, and thus constraint (22b) is equivalent to $z \geq z_2(e,s)$. Using the definition of $W(e,s,z)$ in (12), we obtain that $W(e,s,z) - W^0 = \frac{(\gamma_2+e+s+z)^2-\gamma_2^2}{2\gamma_1} - e$, and thus constraint (22c) is equivalent to $z \geq z_3(e,s)$. Using the definition of $\Pi_s(e,s,z)$ in (15), we obtain that $\Pi_s(e,s,z) - \Pi_s^0 = \frac{(K-z)(e+s+z)}{\gamma_1} - \frac{\gamma_2 z}{\gamma_1}$, and thus constraint (22d) is equivalent to $z_5'(e,s) \leq z \leq z_5(e,s)$, where $z_5'(e,s) = -\sqrt{\left(\frac{K}{2} - \frac{1}{2}\left(\gamma_2 + (e+s)\right)\right)^2 + K(e+s)} + \frac{K}{2} - \frac{1}{2}\left(\gamma_2 + (e+s)\right) \leq 0$. Thus, constraints (22b), (22c) and (22d) along with constraints $z \geq 0$ and $e+s+z \leq \gamma_1 - \gamma_2$ are equivalent to

$$z \in [\max\{z_2(e,s), z_3(e,s), 0\}, \min\{z_4(e,s), z_5(e,s)\}]. \tag{48}$$

For constraint (22e), note that when fixing e and s, $\Pi_s(e,s,z) = (K-z)\frac{e+s+z+\gamma_2}{\gamma_1}$ is concave in z. By solving the first order condition $\frac{\partial \Pi_s(e,s,z)}{\partial z} = 0$, and denoting the solution as $z_1(e,s)$, we obtain $z_1(e,s) = \frac{1}{2}(K - (\gamma_2 + (e+s)))$. It is straightforward to show that $z_1(e,s) \leq z_5(e,s)$. Then for any (e,s,z) that satisfies (48), constraint (22e) is equivalent to $z = \min\{\max\{z_1(e,s), z_2(e,s), z_3(e,s), 0\}, z_4(e,s), z_5(e,s)\}$. Therefore, if $e+s+z \leq \gamma_1 - \gamma_2$, the feasible region of (22) is $\mathcal{P}_1 = \{(e,s,z) |\max\{z_2(e,s), z_3(e,s), 0\} \leq z \leq \min\{z_4(e,s), z_5(e,s)\}, z = \min\{\max\{z_1(e,s), z_2(e,s), z_3(e,s), 0\}, z_4(e,s), z_5(e,s)\}, e, s, z \geq 0\}$.

Next, from the definitions of $z_i(e,s), i \in \{1, 2, \ldots, 5\}$ in Lemma 2, it is straightforward to show that the feasible region of problem (47) is $\mathcal{P}_2 = \{(e,s,z) |\max\{z_2(e,s), z_3(e,s), 0\} \leq \min\{z_4(e,s), z_5(e,s)\}, z = \min\{\max\{z_1(e,s), z_2(e,s), z_3(e,s), 0\}, z_4(e,s), z_5(e,s)\}, e, s, z \geq 0\}$.

It's plain to show that for any $(e', s', z') \in \mathcal{P}_1$, it holds that $(e', s', z') \in \mathcal{P}_2$, and for any $(e'', s'', z'') \in \mathcal{P}_2$, we have $(e'', s'', z'') \in \mathcal{P}_1$. It follows that $\mathcal{P}_1$ is equivalent to $\mathcal{P}_2$. The equivalence between the two problems then follows. $\square$

For use in Lemma 3 below, let $\breve{x} = \frac{(d-p)-(\delta-1)+2K}{(r-d)+(\delta-1)}$, $\bar{\beta}(x) = \frac{2K}{\gamma_2}$, and $\beta_1(x) = \frac{2\gamma_1 - 3\gamma_2 - \sqrt{(2\gamma_1-\gamma_2)^2 - 8\gamma_1 K}}{2\gamma_2}$.

Lemma 3. *Suppose $K < \gamma_1$ and $\beta \in (\tau(x), \bar{\tau}(x))$. Further, suppose the constraint $D_m(e,s,z) = \frac{\gamma_2+e+s+z}{\gamma_1}$, or equivalently, $e+s+z \leq \gamma_1 - \gamma_2$, is added to formulation (22). Then the following statements hold:*

(i) If $x \leq \breve{x}$, then $e^* > 0$, $s^* \geq 0$, and $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$.

(ii) If $x > \breve{x}$, and

(a) if $\tau(x) < \beta \leq \beta_1(x)$, then $e^* > 0, s^* \geq 0$ and $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$.

(b) if $\beta_1(x) < \beta \leq \bar{\beta}(x)$, then $e^* > 0, s^* > 0$, and $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$.

(c) if $\bar{\beta}(x) < \beta < \bar{\tau}(x)$, then $e^* > 0, s^* > 0$, and $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ > 0$.

Further, the optimal objective value is $[C(e^*, s^*, z^*)]^+ = \max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1}\right\}$.

Proof. Note that $\bar{\tau}(x) = \frac{\gamma_1}{\gamma_2} - 1$ and $\tau(x) = \min\left\{\frac{K}{\gamma_2} - 1, \frac{\gamma_1}{\gamma_2} - 1\right\}$. Therefore, $K < \gamma_1$ implies that $(\tau(x), \bar{\tau}(x))$ is non-empty, and that $\tau(x) = \frac{K}{\gamma_2} - 1$. Note that in Lemma 2, we showed that when imposing the additional constraint $e+s+z \leq \gamma_1 - \gamma_2$, solving problem (22) is equivalent to solving problem (47); therefore, we will identify the optimal solution (e^*, s^*, z^*) to (22) by solving (47). The proof proceeds in three steps. In Step 1, we show that any feasible z to problem (47) must satisfy one of the following five equations: $z = z_1(e,s), z = z_2(e,s), z = 0, z = z_3(e,s), z = z_4(e,s)$. In Step 2, we solve for the optimal solutions to problem (47) by imposing each of these five equations as an additional constraint. In Step 3, we identify the optimal solution (e^*, s^*, z^*) in formulation (47) by comparing the optimal objective function values for each of these five subproblems.

Step 1. To solve formulation (47), let us first consider constraint (47b), which requires $z = \breve{z}(e,s)$, where $\breve{z}(e,s) = \min\{\max\{z_1(e,s), z_2(e,s), z_3(e,s), 0\}, z_4(e,s), z_5(e,s)\}$ for simplicity. Note that $\breve{z}(e,s) \neq z_5(e,s)$. To see why this holds, note that from the proof of Lemma 2, $z_1(e,s) < z_5(e,s)$; in addition, constraints (47c) – (47f) are equivalent to $\max\{z_2(e,s), z_3(e,s), 0\} \leq \min\{z_4(e,s), z_5(e,s)\}$. Therefore, $\breve{z}(e,s)$ must satisfy one of the following five conditions: $\breve{z}(e,s) = z_1(e,s), \breve{z}(e,s) = z_2(e,s), \breve{z}(e,s) = 0, \breve{z}(e,s) = z_3(e,s)$, or $\breve{z}(e,s) = z_4(e,s)$.

Step 2. We include each one of the five conditions listed above as additional constraint in problem (47), and solve the corresponding subproblem.

Case 1: Suppose $\breve{z}(e,s) = z_1(e,s)$. We show this case is infeasible. Specifically, if $\breve{z}(e,s) = z_1(e,s)$, then

$$z_1(e,s) \geq z_2(e,s) \text{ if and only if } e+s \geq 2(\beta+1)\gamma_2 - (\gamma_2 + K), \tag{49a}$$

$$z_1(e,s) \geq z_3(e,s) \text{ if and only if } e+s \geq 2\sqrt{\gamma_2^2 + 2\gamma_1 e} - (\gamma_2 + K), \tag{49b}$$

$$z_1(e,s) \geq 0 \text{ if and only if } e+s \leq K - \gamma_2, \tag{49c}$$

$$z_1(e,s) \leq z_4(e,s) \text{ if and only if } e+s \leq 2\gamma_1 - \gamma_2 - K, \tag{49d}$$

and it is trivial to prove $z_1(e,s) \leq z_5(e,s)$. Define $y_1 \equiv \frac{\gamma_2^2 + 2\gamma_1 e + \gamma_2 K}{\sqrt{\gamma_2{}^2 + 2\gamma_1 e}} - (\gamma_2 + K)$, $y_2 \equiv 2(\beta+1)\gamma_2 - (\gamma_2 + K)$, $y_3 \equiv 2\sqrt{\gamma_2{}^2 + 2\gamma_1 e} - (\gamma_2 + K)$ and $y_4 \equiv \frac{\beta((\beta+1)\gamma_2 - K)}{\beta+1}$. Combining constraints (49) with problem (47), then a feasible (e,s) must satisfy

$$e \leq \frac{\gamma_1^2 - \gamma_2^2}{2\gamma_1}, \max\{e, y_1, y_2, y_3\} \leq e+s \leq K - \gamma_2. \tag{50}$$

Because $\beta > \frac{K}{\gamma_2} - 1$, it follows that $\max\{e, y_1, y_2, y_3\} \geq y_2 = 2(\beta+1)\gamma_2 - (\gamma_2 + K) > K - \gamma_2$, which contradicts (50). Therefore, a feasible solution(e,s,z) to problem (22) never satisfies $z = \breve{z}(e,s) = z_1(e,s)$.

Case 2: Suppose $\breve{z}(e,s) = z_2(e,s)$. Similar to Case 1, the condition implies $s+e \leq 2\beta\gamma_2 + \gamma_2 - K$, $e \leq \frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1}$ and $s+e \leq \beta\gamma_2$. Note that $e \leq \frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1}$ implies $\max\{y_1, y_4\} = y_4 < \beta\gamma_2$, where y_1 and y_4 are defined in Case 1. By plugging $z = z_2(e,s)$ into the objective, we obtain the cost function as $C(e,s,z_2(e,s)) = \frac{(\beta+1)\gamma_2}{\gamma_1}(e+s) - e$. Then given the additional constraint $\breve{z}(e,s) = z_2(e,s)$, problem (47) is equivalent to the following linear program:

$$\min_{y \geq e \geq 0} \left[\frac{(\beta+1)\gamma_2}{\gamma_1}(e+s) - e\right]^+ \tag{51a}$$

$$\text{s.t. } e \leq \frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1} \tag{51b}$$

$$\max\{y_4, e\} \leq e + s \leq \beta\gamma_2, \tag{51c}$$

By solving the above linear program, it follows that if $\frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1} \geq y_4$, the optimal solution (e^*, s^*, z^*) must satisfy $e^* > 0$, $s^* \geq 0$, $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$; if $\frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1} < y_4$ and $\beta \leq \frac{2K}{\gamma_2}$, then $e^* > 0$, $s^* > 0$, $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$; if $\frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1} < y_4$ and $\beta > \frac{2K}{\gamma_2}$, then $e^* > 0$,$s^* > 0$,$z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1} > 0$. Note that $\frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1} < y_4$ if and only if $(\gamma_1, \beta) \in \left\{\gamma_1, \beta \mid \frac{\gamma_2}{2} + K + \sqrt{K(\gamma_2 + K)} < \gamma_1 \leq 2K + \gamma_2, \beta \in (\beta_1, \beta_2)\right\} \cup \{\gamma_1, \beta \mid \gamma_1 > 2K + \gamma_2, \beta \in (\beta_1, \bar{\tau}(x))\}$, where $\beta_2 = \frac{2\gamma_1 - 3\gamma_2 + \sqrt{(2\gamma_1 - \gamma_2)^2 - 8\gamma_1 K}}{2\gamma_2}$.

Case 3: Suppose $\check{z}(e, s) = z_3(e, s)$. Using a similar argument to Case 2, by letting $Y = (e+s)\sqrt{\gamma_2^2 + 2\gamma_1 e}$ and $t = \sqrt{\gamma_2^2 + 2\gamma_1 e}$, and rewriting the cost function as as $C_3(Y, t) = C(e, s, z_3(e, s)) = \frac{2Y - t^2 + \gamma_2^2}{2\gamma_1}$, we can show that with the additional constraint $\check{z}(e, s) = z_3(e, s)$, problem (47) is equivalent to the following:

$$\min_{Y, t \geq 0} \ [C_3(Y, t)]^+ \tag{52a}$$

$$\text{s.t. } (\beta+1)\gamma_2 \leq t \leq \gamma_1 \tag{52b}$$

$$\max\left\{t^2 - (\gamma_2 + K)\,t + K\gamma_2, \frac{t\left(t^2 - \gamma_2^2\right)}{2\gamma_1}\right\} \leq Y \leq t^2 - \gamma_2 t \tag{52c}$$

Let $Y_1(t) = t^2 - (\gamma_2 + K)\,t + K\gamma_2$ and $Y_2(t) = \frac{t\left(t^2 - \gamma_2^2\right)}{2\gamma_1}$ for simplicity. Then constraint (52c) becomes $\{Y_1(t), Y_2(t)\} \leq Y \leq t^2 - \gamma_2 t$. Next, we consider two cases according to constraint (52c): $Y_1(t) \leq Y_2(t)$ and $Y_1(t) > Y_2(t)$. In preparation, let $\mathcal{B}_1 = \{\gamma_1, \beta \mid \gamma_1 \leq 2K + \gamma_2\} \cup \left\{\gamma_1, \beta \mid \gamma_1 > 2K + \gamma_2, \beta \in \left(\frac{K}{\gamma_2} - 1, \beta_1\right]\right\}$ and $\mathcal{B}_2 = \{\gamma_1, \beta \mid \gamma_1 > 2K + \gamma_2, \beta \in (\beta_1, \bar{\tau}(x))\}$; then we have if $(\gamma_1, \beta) \in \mathcal{B}_1$, there exist a t satisfying constraint (52b) such that $Y_1(t) \leq Y_2(t)$; if $(\gamma_1, \beta) \in \mathcal{B}_2$, for all t satisfying constraint (52b), it holds that $Y_1(t) > Y_2(t)$.

First, suppose $Y_1(t) \leq Y_2(t)$. We only focus on the regime where $(\gamma_1, \beta) \in \mathcal{B}_1$. It is straightforward to verify that in this case, $e^* \geq 0$, $s^* \geq 0$, $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$. Next, when restricting $Y_1(t) > Y_2(t)$, note that we have already proved that when $(\gamma_1, \beta) \in \mathcal{B}_1$, (e^*, s^*, z^*) satisfy $e^* > 0$, $s^* \geq 0$ and $z^* > 0$, so we only need to consider the regime where $(\gamma_1, \beta) \in \mathcal{B}_2$. In this case, it is straightforward to show that if $2K + \gamma_2 \geq (\beta+1)\gamma_2$, or equivalently $\beta \leq \frac{2K}{\gamma_2}$, then $e^*, s^*, z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$; if $2K + \gamma_2 < (\beta+1)\gamma_2$, or equivalently $\beta > \frac{2K}{\gamma_2}$, then $e^*, s^*, z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1} > 0$.

Case 4: Suppose $\check{z}(e, s) = 0$. We show that this case is dominated by Case 3. Specifically, when restricting $\check{z}(e, s) = 0$, it is easy to show that problem (47) reduces to (16). Thus the minimal cost in this case must be no less than that of mechanism [D], namely, $C(e^*, s^*, z^*) \geq \frac{\beta^2\gamma_2^2}{2\gamma_1} > 0$. Comparing the minimal objective value in this case with that in Case 3, which is given by $\max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1}\right\}$, we can show that $[C(e^*, s^*, z^*)]^+ = C(e^*, s^*, z^*) \geq \frac{\beta^2\gamma_2^2}{2\gamma_1} > \max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1}\right\}$, which implies this case is dominated by Case 3.

Case 5: Suppose $\check{z}(e, s) = z_4(e, s)$. We show this case is infeasible. Specifically, assuming $\check{z}(e, s) =$

$z_4(e,s)$, this is equivalent to

$$z_1(e,s) \geq z_2(e,s) \text{ if and only if } e+s \geq y_2 = 2(\beta+1)\gamma_2 - (\gamma_2+K), \tag{53a}$$

$$z_1(e,s) \geq z_3(e,s) \text{ if and only if } e+s \geq y_3 = 2\sqrt{\gamma_2^2 + 2\gamma_1 e} - (\gamma_2+K), \tag{53b}$$

$$z_1(e,s) \geq 0 \text{ if and only if } e+s \leq K - \gamma_2, \tag{53c}$$

$$z_1(e,s) \geq z_4(e,s) \text{ if and only if } e+s \geq 2\gamma_1 - \gamma_2 - K, \tag{53d}$$

and $\max\{z_2(e,s), z_3(e,s), 0\} \leq \min\{z_4(e,s), z_5(e,s)\}$. Combining with constraints (47b)–(47f), because $\gamma_1 > K$ and thus $2\gamma_1 - \gamma_2 - K > \gamma_1 - \gamma_2$, inequality (53d) contradicts constraint (47e), which requires $e + s \leq \gamma_1 - \gamma_2$. Therefore, $z = \breve{z}(e,s) = z_4(e,s)$ cannot hold.

Step 3. We have established in Step 2 that the optimal private subsidy z must be equal to either $z_2(e,s)$ or $z_3(e,s)$. Combining Case 2 and Case 3, we conclude that if $(\gamma_1, \beta) \in \{\gamma_1, \beta \mid \gamma_1 \leq 2K + \gamma_2\} \cup \left\{\gamma_1, \beta \mid \gamma_1 > 2K + \gamma_2, \beta \in \left(\frac{K}{\gamma_2} - 1, \beta_1\right]\right\}$, the optimal solution (e^*, s^*, z^*) must satisfy $e^* > 0$, $s^* \geq 0$ and $z^* > 0$; if $(\gamma_1, \beta) \in \{\gamma_1, \beta \mid \gamma_1 > 2K + \gamma_2, \quad \beta \in (\beta_1, \bar{\tau}(x))\}$, the optimal solution (e^*, s^*, z^*) must satisfy $e^*, s^*, z^* > 0$. Moreover, $[C(e^*, s^*, z^*)]^+ = \max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1}\right\}$. Note that by plugging in $\gamma_1 = (\delta-1)(x+1)$ and $\gamma_2 = (d-p) - x(r-d)$, $\gamma_1 \leq 2K + \gamma_2$ is equivalent to $x \leq \breve{x}$, where $\breve{x} = \frac{(d-p)-(\delta-1)+2K}{(r-d)+(\delta-1)}$. The result follows. □

Proof of Proposition 5. Similar to the proofs of Propositions 1 and 2, the proof is composed of two cases: $D_m(e,s,z) = \frac{\gamma_2+e+s+z}{\gamma_1}$, or equivalently, $e+s+z \leq \gamma_1 - \gamma_2$; and $D_m(e,s,z) = 1$, or equivalently, $e+s+z \geq \gamma_1 - \gamma_2$.

Case 1: Suppose (e,s,z) is restricted to the regime where $D_m(e,s,z) = \frac{\gamma_2+e+s+z}{\gamma_1}$, which is equivalent to $e+s+z \leq \gamma_1 - \gamma_2$. We divide the analysis into two further cases with respect to β: $\beta \in (0, \tau(x)]$ and $\beta \in (\tau(x), \bar{\tau}(x))$.

Case 1.1: Suppose $\beta \in (0, \tau(x)]$. Because $[C(e,s,z)]^+ \geq 0$, if a feasible solution (e,s,z) satisfies $C(e,s,z) \leq 0$, it must be optimal. Note that the minimal cost of mechanism [I] is 0, so the optimal solution to problem (18) under mechanism [I] is also optimal to (22) in the case when $\beta \in (0, \tau(x)]$. Namely, there exists an optimal solution to (22) $(e^*, s^*, z^*) = (0, 0, \min\{\max\{\underline{z}, \hat{z}\}, \bar{z}\})$, where $\underline{z}, \hat{z}, \bar{z}$ are as defined in Proposition 2. It can also be shown that $(e^*, s^*, z^*) \in \mathcal{Q}$ is optimal, where $\mathcal{Q}$ is defined as $\mathcal{Q} = \left\{(e^*, s^*, z^*) | e^* = \frac{\beta(\beta+2)\gamma_2^2}{2\gamma_1}, \;\; s^* \in \left[0, \frac{\beta(\beta+2)\gamma_2(\gamma_1 - (\beta+1)\gamma_2)}{2(\beta+1)\gamma_1}\right], z^* = z_2(e^*, s^*)\right\}$, and the objective value is $[C(e^*, s^*, z^*)]^+ = 0$. Note that the optimal private subsidy must be strictly positive, $z^* > 0$; otherwise, if $z^* = 0$, the optimal cost should be no less than that of mechanism [D] as presented in Corollary 2 and is thus strictly positive. To summarize, when $\beta \in (0, \tau(x)]$, the optimal solutions (e^*, s^*, z^*) must satisfy $e^*, s^* \geq 0$ and $z^* > 0$ and $[C(e^*, s^*, z^*)]^+ = 0$.

Case 1.2: When $\beta \in (\tau(x), \bar{\tau}(x))$, the optimal solutions have been characterized in Lemma 3.

To summarize, when requiring $D_m(e,s,z) = \frac{\gamma_2+e+s+z}{\gamma_1}$, the optimal objective value is $[C(e^*, s^*, z^*)]^+ = \max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1}\right\}$ and the optimal solution (e^*, s^*, z^*) is characterized by the following:

(i) Suppose $x \leq \breve{x}$. Then $[C(e^*, s^*, z^*)]^+ = 0$. Further,

(a) if $0 \le \beta \le \tau(x)$, then $e^* \ge 0, s^* \ge 0$ and $z^* > 0$;

(b) if $\tau(x) < \beta < \bar{\tau}(x)$, then $e^* > 0, s^* \ge 0$, and $z^* > 0$.

(ii) Suppose $x > \breve{x}$. Then

(a) if $0 \le \beta \le \bar{\beta}(x)$, $[C(e^*, s^*, z^*)]^+ = 0$. Further,

(1) if $0 \le \beta \le \tau(x)$, then $e^* \ge 0, s^* \ge 0$, and $z^* > 0$;

(2) if $\tau(x) < \beta \le \beta_1(x)$, then $e^* > 0, s^* \ge 0$ and $z^* > 0$;

(3) if $\beta_1(x) < \beta \le \bar{\beta}(x)$, then $e^* > 0, s^* > 0$, and $z^* > 0$

(b) if $\bar{\beta}(x) < \beta < \bar{\tau}(x)$, $[C(e^*, s^*, z^*)]^+ = \frac{\beta\gamma_2(\beta\gamma_2 - 2K)}{2\gamma_1} > 0$. Further, $e^* > 0, s^* > 0$, and $z^* > 0$.

Case 2: Suppose $D_m(e, s, z) = 1$, which is equivalent to $e + s + z \ge \gamma_1 - \gamma_2$. We show this case is never optimal. Using the definition in (15), in this case, $\Pi_s(e, s, z) = (K - z) \cdot D_m(e, s, z) = K - z$. Then from constraint (22e), a feasible private subsidy z must satisfy $z = \arg\max_z \Pi_s(e, s, z) = \arg\max_z (K - z) = 0$. Using the definition in (12), constraint (22c) can then be rewritten as

$$W(e, s, z) - W^0 = s + z - \frac{(\gamma_1 - \gamma_2)^2}{2\gamma_1} = s - \frac{(\gamma_1 - \gamma_2)^2}{2\gamma_1} \ge 0, \tag{54}$$

which implies feasible s must satisfy $s \ge \frac{(\gamma_1-\gamma_2)^2}{2\gamma_1}$. Recall from the proof of Proposition 1, when $D_m(e, s, z) = 1$, $C(e, s, z) = s \ge \frac{(\gamma_1-\gamma_2)^2}{2\gamma_1} > \frac{\beta^2\gamma_2^2}{2\gamma_1}$, where $\frac{\beta^2\gamma_2^2}{2\gamma_1}$ is the minimal cost under mechanism [D]. Further, because the optimal objective value in Case 1 is $\max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2-2K)}{2\gamma_1}\right\}$ and it is straightforward to verify that the $\frac{\beta^2\gamma_2^2}{2\gamma_1} > \max\left\{0, \frac{\beta\gamma_2(\beta\gamma_2-2K)}{2\gamma_1}\right\}$, it follows the optimal solution (e^*, s^*, z^*) must satisfy $D_m(e, s, z) = \frac{\gamma_2+e+s+z}{\gamma_1}$. Hence $D_m(e^*, s^*, z^*) = 1$ cannot hold at optimality. □

Proof of Corollary 5. Using the optimal solutions to mechanism [G] presented in the proof of Proposition 5, it follows the optimal objective value under mechanism [G] is

$$[C(e^*, s^*, z^*)]^+ = \begin{cases} 0 & \text{if } \beta \le \bar{\beta}(x), \\ \frac{\beta\gamma_2(\beta\gamma_2-2K)}{2\gamma_1} & \text{if } \beta > \bar{\beta}(x), \end{cases}$$

where $\bar{\beta}(x) = \frac{2K}{\gamma_2}$ and γ_1, γ_2 are defined as in Proposition 5. Using the definition above, it is straightforward to show that $[C(e^*, s^*, z^*)]^+$ increases in β, and is weakly less than the minimal costs of both mechanisms [D] and [I]. □

Proof of Proposition 6. The proof consists of two steps. In Step 1, we establish an equivalent formulation to problem (23). In Step 2, we solve for the optimal solutions to the equivalent problem and problem (23).

Step 1. We first prove problem (23) is equivalent to the following problem:

$$\max_{e,s\ge 0} \min\{e + s, \gamma_1 - \gamma_2\} \tag{55a}$$

$$\text{s.t. } C(e, s, 0) \le B \tag{55b}$$

$$W(e, s, 0) \ge W^0. \tag{55c}$$

To show the equivalence between the two problems, first note that from the definition of $D_m(e,s,z)$ in (10), we have $D_m(e,s,0)-D_m^0=\frac{\min\{e+s,\gamma_1-\gamma_2\}}{\gamma_1}$, where $\gamma_1=(\delta-1)(x+1)$ and $\gamma_2=(d-p)-x(r-d)$. Because $\delta>1$ and $x>0$, maximizing $D_m(e,s,0)-D_m^0$ is equivalent to maximizing $\min\{e+s,\gamma_1-\gamma_2\}$. Further, from the definition of $\Pi_s(e,s,z)$ in (15), $\Pi_s(e,s,z)-\Pi_s^0=(K-z)\frac{\min\{e+s,\gamma_1-\gamma_2\}}{\gamma_1}$. Because $\underline{x}<x<\bar{x}$, then $\gamma_1-\gamma_2>0$. Therefore, $e,s\geq 0$ implies $\Pi_s(e,s,z)\geq\Pi_s^0$, which allows us to omit the corresponding constraint in (23) without changing the feasible region. Therefore, problem (23) is equivalent to problem (55).

Step 2. To solve problem (55), we change variables by letting $y=e+s$. Then using the definitions of $C(e,s,0)$ in (11) and $W(e,s,0)$ in (12), formulation (55) can be rewritten as

$$\max_{0\leq s\leq y}\ \min\{y,\gamma_1-\gamma_2\} \tag{56a}$$

$$\text{s.t. } s\leq B+\frac{\left(\frac{\gamma_1-\gamma_2}{2}\right)^2-\left(\min\{y,\gamma_1-\gamma_2\}-\frac{\gamma_1-\gamma_2}{2}\right)^2}{\gamma_1} \tag{56b}$$

$$s\geq\frac{(\gamma_1-\gamma_2)^2-(\min\{y,\gamma_1-\gamma_2\}-(\gamma_1-\gamma_2))^2}{2\gamma_1}. \tag{56c}$$

Let us relax the non-negativity constraint $0\leq s\leq y$ for now. Note that a feasible y must satisfy

$$\frac{(\gamma_1-\gamma_2)^2-(\min\{y,\gamma_1-\gamma_2\}-(\gamma_1-\gamma_2))^2}{2\gamma_1}\leq B+\frac{\left(\frac{\gamma_1-\gamma_2}{2}\right)^2-\left(\min\{y,\gamma_1-\gamma_2\}-\frac{\gamma_1-\gamma_2}{2}\right)^2}{\gamma_1}, \tag{57}$$

or equivalently, $B\geq\frac{(\min\{y,\gamma_1-\gamma_2\})^2}{2\gamma_1}$. Define $\tilde{B}=\frac{(\gamma_1-\gamma_2)^2}{2\gamma_1}$. We then decompose the analysis in two cases: (i) when $B\leq\tilde{B}$ and (ii) when $B>\tilde{B}$, which corresponds to statements (i) and (ii) in the proposition.

- (i) If $B\leq\tilde{B}$, a feasible y must satisfy $\frac{(\min\{y,\gamma_1-\gamma_2\})^2}{2\gamma_1}\leq\tilde{B}=\frac{(\gamma_1-\gamma_2)^2}{2\gamma_1}$, which implies $y\leq\gamma_1-\gamma_2$, and in this case, (57) is equivalent to $y^2\leq 2\gamma_1 B$. Because $\min\{y,\gamma_1-\gamma_2\}=y$, then the unique optimal solution is $y^*=\sqrt{2\gamma_1 B}\leq\gamma_1-\gamma_2$ and $s^*=B\left(\frac{\sqrt{2}(\gamma_1-\gamma_2)}{\sqrt{\gamma_1 B}}-1\right)$. It follows that $e^*=y^*-s^*=\frac{\sqrt{2B}\gamma_2}{\sqrt{\gamma_1}}+B$ and $D_m(e^*,s^*,0)=\frac{\sqrt{2B\gamma_1}+\gamma_2}{\gamma_1}$. Because $y^*\geq s^*\geq 0$, then (y^*,s^*) is optimal to problem (56) and thus (e^*,s^*) is optimal to both problems (23) and (55).
- (ii) Suppose $B>\tilde{B}$. We consider two subproblems, that follow from imposing the additional constraints $y\geq\gamma_1-\gamma_2$ and $y<\gamma_1-\gamma_2$. When $y\geq\gamma_1-\gamma_2$, then the optimal solution (s^*,y^*) is such that $s^*\in[\tilde{B},B]$ and $y^*\in[\gamma_1-\gamma_2,\infty)$. The optimal objective value is $\gamma_1-\gamma_2$. In this case, it must hold that $e^*\geq\gamma_1-\gamma_2-s^*$ and $D(e^*,s^*,0)=1$. However, when $y<\gamma_1-\gamma_2$, the optimal objective value is strictly less than $\gamma_1-\gamma_2$, which is dominated by the case when $y\geq\gamma_1-\gamma_2$. Because $0\leq s^*\leq y^*$, (y^*,s^*) is optimal to problem (56).

The result follows by plugging in $\gamma_1=(\delta-1)(x+1)$ and $\gamma_2=(d-p)-x(r-d)$. □

Proof of Proposition 7. Similar to the proof of Proposition 4, it is straightforward to show that constraints (25d) and (25e) uniquely determines a solution $z^*=\min\{\hat{z},(\delta-1)(x+1)-((d-p)-x(r-d))\}$, where $\hat{z}=\frac{1}{2}(K-((d-p)-x(r-d)))$. Using (11) and (12), we find z^* satisfies constraints (25b) and (25c). Therefore, z^* is the unique feasible and thus optimal solution to problem (25). □

Proof of Proposition 8. Similar to the proof of Proposition 1, to solve for the optimal solution (e^*,s^*), we consider two cases using the definition of $D_m(e,s,0)$ in (27), which implies $D_m(e,s,0) = \min\left\{\sigma, \frac{(d-p)-(\delta-1)V+e+s}{(r-d)+(\delta-1)V}\right\}$: $D_m(e,s,0) = \frac{(d-p)-(\delta-1)V+e+s}{(r-d)+(\delta-1)V}$ and $D_m(e,s,0) = \sigma$.

Case 1: Suppose we add $D_m(e,s,0) = \frac{(d-p)-(\delta-1)V+e+s}{(r-d)+(\delta-1)V}$ as an additional constraint. Using a similar argument to the proof of Lemma 1, it can be shown that

$$[C(e,s,0)]^+ \geq C(e,s,0) \geq \frac{\beta^2((d-p)-(\delta-1)V)^2}{2((r-d)+(\delta-1)V)},$$

which holds with equality if and only if both constraints (29b) and (29c) are binding. To find the optimal solution, it suffices to construct a feasible solution (e_1, s_1) that attains the lower bound, where (e_1, s_1) is the unique solution to the equations

$$\begin{aligned} D_m(e,s,0) - D_m^0 &= \beta D_m^0 \\ W(e,s,0) &= W^0, \end{aligned}$$

which correspond to constraints (29b) and (29c). Then (e_1, s_1) is given by:

$$\begin{aligned} e_1 &= \frac{\beta(\beta+2)((d-p)-(\delta-1)V)^2}{2\sigma((r-d)+(\delta-1)V)}, \\ s_1 &= \beta((d-p)-(\delta-1)V) - \frac{\beta(\beta+2)((d-p)-(\delta-1)V)^2}{2\sigma((r-d)+(\delta-1)V)}. \end{aligned}$$

Using a similar argument to the proof of Proposition 1, it can be verified that (e_1, s_1) satisfies $D_m(e,s,0) = \frac{(d-p)-(\delta-1)V+e+s}{(r-d)+(\delta-1)V}$ and is feasible to (29). Therefore, (e_1, s_1) is the unique optimal solution to problem (29) when requiring $D_m(e,s,0) = \frac{(d-p)-(\delta-1)V+e+s}{(r-d)+(\delta-1)V}$.

Case 2: Suppose $D_m(e,s,0) = \sigma$. It follows that $D_d(e,s,0) = 0$ and the operational cost is $C(e,s,0) = s\sigma$. In this case, constraint (29c) can be rewritten as

$$W(e,s,0) - W^0 = s\sigma - \frac{((d-p)-\sigma((r-d)+(\delta-1)V)-(\delta-1)V)^2}{2((r-d)+(\delta-1)V)} \geq 0. \tag{58}$$

Let $\phi \equiv \frac{((d-p)-\sigma((r-d)+(\delta-1)V)-(\delta-1)V)^2}{2((r-d)+(\delta-1)V)}$. Then constraint (29c) implies $s\sigma \geq \phi \geq 0$ and thus that $[C(e,s,0)]^+ = C(e,s,0) = s\sigma \geq \phi$. It is straightforward to verify that $\phi > C(e_1, s_1, 0) = [C(e_1, s_1, 0)]^+$, where (e_1, s_1) is the optimal solution in Case 1, and it follows that $[C(e,s,0)]^+ \geq \phi > [C(e_1, s_1, 0)]^+$. Therefore, the optimal solution to (29) is $(e^*, s^*) = (e_1, s_1)$, and the resulting minimal cost is strictly positive, $C(e_1, s_1, 0) = \frac{\beta^2((d-p)-(\delta-1)V)^2}{2((r-d)+(\delta-1)V)}$. □

Proof of Proposition 9. Using a parallel argument to the proof of Proposition 2, we first note that $[C(0,0,z)]^+ = C(0,0,z) = 0$. As such, any feasible solution to problem (29) is optimal. Therefore, to find the feasible and thus optimal solution to problem (29), it suffices to find the optimal solution z^* to the following problem:

$$\max_{z\geq 0} \Pi_s(0,0,z) \tag{59a}$$
$$\text{s.t. } D_m(0,0,z) - D_m^0 \geq \beta D_m^0, \tag{59b}$$
$$W(0,0,z) \geq W^0, \tag{59c}$$
$$\Pi_s(0,0,z) \geq \Pi_s^0. \tag{59d}$$

Using a similar argument in the proof of Proposition 2, it is straightforward to show that the feasible region of (59) is $[\underline{z},\bar{z}]\cup(\tilde{z},\dot{z}]$, where $\underline{z}=\beta((d-p)-(\delta-1)V)$, $\bar{z}=\min\{K,\sigma((r-d)+(\delta-1)V)\}-((d-p)-(\delta-1)V)$, $\tilde{z}=\sigma((r-d)+(\delta-1)V)-((d-p)-(\delta-1)V)$ and $\dot{z}=K\left(1-\frac{(d-p)-(\delta-1)V}{\sigma((r-d)+(\delta-1)V)}\right)$. Note (59) and (31) are feasible if and only if $[\underline{z},\bar{z}]\cup(\tilde{z},\dot{z}]$ is non-empty, or equivalently, $\beta\leq\tau(x)$, where $\tau(x)=\min\left\{\frac{K}{(d-p)-(\delta-1)V}-1,\bar{\tau}(x)\right\}$ and $\bar{\tau}(x)=\frac{\sigma((r-d)+(\delta-1)V)}{(d-p)-(\delta-1)V}-1$. We can now solve (59) assuming $\beta\leq\tau(x)$. Similar to the proof of Proposition 2, we can show it suffices to solve the following problem

$$\max \Pi_s(0,0,z) \text{ s.t. } z\in[\underline{z},\bar{z}].$$

Because $\frac{\partial^2\Pi_s(0,0,z)}{\partial z^2}=-\frac{2}{(r-d)+(\delta-1)V}<0$, it follows that $\Pi_s(0,0,z)$ is strictly concave in z. The solution to the first order condition $\frac{\partial\Pi_s(0,0,z)}{\partial z}=0$ is $\hat{z}=\frac{K}{2}-\frac{1}{2}((r-d)+(\delta-1)V)$. Therefore, the unique optimal subsidy z^* is given by

$$z^*=\begin{cases}\underline{z} & \text{if } \hat{z}<\underline{z},\\ \hat{z} & \text{if } \underline{z}\leq\hat{z}\leq\bar{z},\\ \bar{z} & \text{if } \bar{z}<\hat{z},\end{cases}$$

or, equivalently, $z^*=\min\{\max\{\underline{z},\hat{z}\},\bar{z}\}$. It follows that z^* is the unique feasible and thus optimal solution to (31). The result follows. □

Proof of Theorem 3. Similar to the proof of Theorem 1, using the optimal incentives under mechanisms [D-H] and [D-I] presented in Propositions 8 and 9, it is straightforward to show that $C^I=0<C^D$ and $W^I>W^0=W^D$. Also, it can be shown that $D_m(0,0,z^*)-D_m^0\geq\beta D_m^0=D_m(e^*,s^*,0)-D_m^0$, or equivalently, $D_m^I\geq D_m^D$. Using the definitions in (13) and (14), it then follows that $\Pi_p^I\geq\Pi_p^D$ and $\Pi_r^I\geq\Pi_r^D$, respectively. With respect to the enterprise's profit, from (15) we have $\Pi^I-\Pi^D=\frac{(K-z^*)((d-p)-(\delta-1)V+z^*)}{(r-d)+(\delta-1)V}-\frac{(\beta+1)K((d-p)-(\delta-1)V)}{(r-d)+(\delta-1)V}$. Define $\tilde{\tau}(x)=\frac{\hat{z}^2-(\hat{z}-z^*)^2}{K((d-p)-(\delta-1)V)}$. Then $\Pi_s^I<\Pi_s^D$ holds if and only if $\beta>\tilde{\tau}(x)$. The results follow. □