A. Notational glossary

Table A.5 provides a glossary of the mathematical symbols used.

Table A.5: Notational glossary.

Symbol	**Name**
N	Number of commuters
s	Bottleneck capacity
s^{fast}	- fast lane
s^{slow}	- slow lane
T	Number of discrete time intervales
Δ	Length of discrete time interval
τ	Commuter type
g_τ	- distribution of types
$[u, k]$	Commuter state
u	- Value of Time (VOT)/urgency
k	- karma
$\bar{k}$	Average karma per commuter
$[t, b]$	Commuter action
t	- departure time
b	- karma bid
b^*	Threshold bid to enter fast lane
t^*_τ	Desired arrival time
ϕ_τ	VOT process
$\bar{u}_\tau$	Long-term average VOT
(d, π)	Social state
d	- distribution of types and states
π	- policy (of all types)
π_τ	- policy of type τ
(d^*, π^*)	Stationary Nash Equilibrium (SNE)
ν	Distribution of actions
ζ_τ	Immediate reward function
κ	Karma transition function
q	Queue length
	Travel delays
t^{q}	- queuing
t^{e}_τ	- early arrival
t^{l}_τ	- late arrival
	Normalized delay penalties
α	- queuing
β	- early arrival
γ	- late arrival

Symbol	**Name**
o	Auction outcome (fast/slow lane)
ψ	Probability of auction outcome
ψ^ϵ	Continuous approximation of ψ
ϵ	- approximation parameter
p	Karma payment
$\bar{p}$	Average payment
f	Fraction of commuters receiving $\lceil \bar{p} \rceil$
$\bar{t}^{\text{q}}_\tau$	Long-term average queuing delay
$\bar{c}_\tau$	Long-term average travel cost
$\bar{c}^{\text{n}}_\tau$	- normalized by $\bar{u}_\tau$
$\bar{t}^{\text{q}}$	System average queuing delay
$\bar{c}^{\text{n}}$	System average normalized travel cost
λ_τ	Income-heterogeneity scaling factor
	Variables for CARMA+
$n[t]$	- mass of commuters departing at t
$r[t]$	- karma redistribution to t
$t^{\text{min-re}}$	- time of minimum redistribution
σ	- ratio of max to min redistribution
$m[t]$	- linear redistribution scaling factor
	Benchmark schemes
NOM	- nominal (no policy intervention)
TOLL	- optimal tolling
c^*	Nominal normalized equilibrium cost
	Nominal bottleneck queue times
t^{start}	- start time
t^{end}	- end time
t^{peak}	- peak time
t^{bnd}	- boundary time in double peak
	Variables for Dynamic Population Game (DPG)
ρ_τ	- state transition function
R_τ	- expected immediate reward
P_τ	- state transition matrix
V_τ	- infinite-horizon reward/value function
Q_τ	- state-action value function
B_τ	- best response
δ	- discount factor

B. Bottleneck model preliminaries

In this section, we review some results from the classical bottleneck model, following the same notations introduced in CARMA. Consider a population of N commuters passing through a single bottleneck with capacity s. Under the optimal tolling scheme, the bottleneck is split into two lanes: a tolled lane with capacity s^{fast} and a regular lane with capacity s^{slow}. Following the setting of CARMA, we assume commuters are heterogeneous in both VOT and desired arrival time, and they can be classified into types. Commuters in the same type $\tau \in \Gamma$ share the same desired arrival time t^*_τ but may differ in their VOT values (due to the dynamic VOT process). Nevertheless, the distribution of VOT values within each type is stationary and exogenous. Besides, all commuters share the same penalty ratios as per Assumption 1.

Hence, the *travel cost* for commuters of type τ with VOT value u who enter the fast lane at time t reads

$$c^{\text{fast}}_\tau[u](t) = \begin{cases} u\,\beta\,(t^*_\tau - t), & t \le t^*_\tau, \\ u\,\gamma\,(t - t^*_\tau), & t > t^*_\tau. \end{cases} \tag{B.1}$$

and their total cost is given by $\hat{c}^{\text{fast}}_\tau[u](t) \coloneqq c^{\text{fast}}_\tau[u](t) + p(t)$, where $p(t)$ is the toll price at time t. For those traveling on the slow lane, the travel cost, as well as the total travel cost, is

$$c^{\text{slow}}_\tau[u](t) = \begin{cases} u\left[\alpha\,\frac{q(t)}{s^{\text{slow}}} + \beta\,(t^*_\tau - t - \frac{q(t)}{s^{\text{slow}}})\right], & t + \frac{q(t)}{s^{\text{slow}}} \le t^*_\tau, \\ u\left[\alpha\,\frac{q(t)}{s^{\text{slow}}} + \gamma\,(t + \frac{q(t)}{s^{\text{slow}}} - t^*_\tau)\right], & t + \frac{q(t)}{s^{\text{slow}}} > t^*_\tau, \end{cases} \tag{B.2}$$

where $q(t)$ denotes the queue length on the slow lane at time t.

The objective of commuters is to minimize their own total travel cost by choosing the lane and departure time. For commuters of type τ and traveling on the slow lane, the equilibrium condition yields the derivative of queue length on the slow lane with respect to departure time t during the congestion period as follows:

$$q'(t) = \begin{cases} \frac{\beta}{\alpha-\beta}\,s^{\text{slow}}, & t + \frac{q(t)}{s^{\text{slow}}} \le t^*_\tau, \\ -\frac{\gamma}{\alpha+\gamma}s^{\text{slow}}, & t + \frac{q(t)}{s^{\text{slow}}} > t^*_\tau. \end{cases} \tag{B.3}$$

Similarly, the derivative of the optimal toll price for the departure window of commuters of type τ and with urgency u on the fast lane is given by

$$p'(t) = \begin{cases} u\beta, & t < t^*_\tau, \\ -u\gamma, & t > t^*_\tau. \end{cases} \tag{B.4}$$

Note that (B.3) applies to `NOM` by replacing s^{slow} with s. These results are used to derive the equilibrium cost and the optimal toll price presented below.

B.1. Homogeneous desired arrival time

In this section, we present the results in the case of homogeneous desired arrival time, i.e., $t^*_\tau \equiv t^*$, for all $\tau \in \Gamma$. Using (B.3), we can easily draw the aggregate departure pattern under `NOM` and derive the start and end of the congestion period (t^{start} and t^{end}), as well as the time with maximum queuing delay (t^{peak}), as follows:

$$t^{\text{start}} = t^* - c^*/\beta, \quad t^{\text{end}} = t^* + c^*/\gamma, \quad t^{\text{peak}} = t^* - c^*/\alpha, \qquad \text{(B.5)} \qquad c^* = \frac{\beta\gamma}{\beta+\gamma}\frac{N}{s}, \qquad \text{(B.6)}$$

where c^* is the VOT-normalized equilibrium cost. For commuters with VOT u, the equilibrium cost is given by $c^*[u] = uc^*$. Another well-known result is that the queuing delay takes half of the total cost [29] and thus $\bar{t}^{\,\text{q}}[u] = \frac{c^*}{2\alpha}$, for all u. These lead to the performance measures listed in the first column of Table 1.

Under TOLL, commuters with higher VOT u would enter the fast lane and depart closer to t^*, and the departure windows are exclusive and allocated on both sides of t^* [19]. Note that the congestion period under TOLL remains the same as NOM, but the system average queuing delay reduces to $\frac{s^{\text{slow}}}{s}\frac{c^*}{2\alpha}$ because the queue only emerges on the slow lane. Let $\hat{u}$ be the threshold VOT such that commuters with $u > \hat{u}$ all enter the fast lane and those with $u < \hat{u}$ all enter the slow lane. Then, the fraction of commuters with $\hat{u}$ who enter the slow lane is given by

$$\hat{r} = \frac{1}{\mathbb{P}[\hat{u}]}\left(\frac{s_{\text{slow}}}{s} - \sum_{u<\hat{u}} \mathbb{P}[u]\right), \qquad \text{(B.7)}$$

where $\mathbb{P}[u]$ gives the fraction of commuters with VOT u. Accordingly, the average queuing delay and travel cost for commuters with VOT u are derived as

$$\bar{t}^{\,\text{q}}[u] = \begin{cases} \frac{c^*}{2\alpha}, & u < \hat{u}, \\ \hat{r}\,\frac{c^*}{2\alpha}, & u = \hat{u}, \\ 0, & u > \hat{u}, \end{cases} \qquad \text{(B.8)} \qquad c^*[u] = \begin{cases} u\,c^*, & u < \hat{u}, \\ u\,c^*\left[1 - (1-\hat{r})\,\frac{s}{2\,s^{\text{fast}}}\,\mathbb{P}[\hat{u}]\right], & u = \hat{u}, \\ u\,c^*\left(1 - \frac{s}{2\,s^{\text{fast}}}\,\mathbb{P}[u]\right), & u > \hat{u}. \end{cases} \qquad \text{(B.9)}$$

(B.8) and (B.9) yield the formulae presented in the second column of Table 1.

B.2. Heterogeneous desired arrival time

We proceed to specify the results for heterogeneous desired arrival times with two types, namely, Type 1 and Type 2 commuters. For notation simplicity, we denote their desired arrival times as t_1^* and t_2^*, respectively. As suggested by [49], the bottleneck may exhibit a single peak or multiple peaks on the slow lane (see, e.g., second right-most column of Figure 8). In what follows, we derive the closed-form performance measures for these two cases, along with the precongestion case considered in Section B.2.

B.2.1. Single peak

When the congestion period exhibits a single peak, the aggregate departure pattern remains the same as in Section B.1. Hence, the system average queuing delay is still $\frac{c^*}{2\alpha}$ under NOM and $\frac{s^{\text{slow}}}{s}\frac{c^*}{2\alpha}$ under TOLL. Besides, the last Type 2 commuter under both NOM and TOLL arrives at t_2^*.

Table B.6: Performance measures with heterogeneous desired arrival time.

Name		Benchmark ("NOM")	Optimal tolling ("TOLL")
Single peak			
System average queuing delay	$\bar{t}^{\,\text{q}}$	$\frac{c^*}{2\alpha}$	$\frac{s^{\text{slow}}}{s}\,\frac{c^*}{2\alpha}$
System average travel cost	$\bar{c}^{\,\text{n}}$	$c^* - \frac{N_2}{N}\beta(t_1^* - t_2^*)$	$\sum_\tau g_\tau\, \bar{c}_\tau^{\text{n}}$
Type average queuing delay	$\bar{t}_1^{\text{q}}$	$\frac{c^*}{2\alpha} + \frac{N_2}{N_1}\frac{\beta}{2\alpha}(t_1^* - t_2^*)$	$\sum_u \mathbb{P}_\tau[u]\, \bar{t}_\tau^{\text{q}}[u]$
	$\bar{t}_2^{\text{q}}$	$\frac{c^*}{2\alpha} - \frac{\beta}{2\alpha}(t_1^* - t_2^*)$	with (B.15)
Type average travel cost	$\bar{c}_1^{\text{n}}$	c^*	$\frac{1}{\bar{u}_\tau}\sum_u \mathbb{P}_\tau[u]\, c_\tau^*[u]$
	$\bar{c}_2^{\text{n}}$	$c^* - \beta(t_1^* - t_2^*)$	with (B.16)
Double peak			
System average queuing delay	$\bar{t}^{\,\text{q}}$	(B.23)	$\frac{s^{\text{slow}}}{s}$ (B.23)
System average travel cost	$\bar{c}^{\,\text{n}}$	(B.20)	$\sum_\tau g_\tau\, \bar{c}_\tau^{\text{n}}$
Type average queuing delay	$\bar{t}_1^{\text{q}}$	(B.21)	$\sum_u \mathbb{P}_\tau[u]\, \bar{t}_\tau^{\text{q}}[u]$
	$\bar{t}_2^{\text{q}}$	(B.22)	with (B.15)
Type average travel cost	$\bar{c}_1^{\text{n}}$	$\frac{1}{2}\left[c^* + \gamma(t_2^* - t_1^* + \frac{N_1}{s})\right]$	$\frac{1}{\bar{u}_\tau}\sum_u \mathbb{P}_\tau[u]\, c_\tau^*[u]$
	$\bar{c}_2^{\text{n}}$	$\frac{1}{2}\left[c^* + \beta(t_2^* - t_1^* + \frac{N_2}{s})\right]$	with (B.16)
Precongestion			
System average queuing delay	$\bar{t}^{\,\text{q}}$	$\frac{c^*}{2\alpha}(\frac{N_1}{N})^2$	$\frac{s^{\text{slow}}}{s}\,\frac{c^*}{2\alpha}(\frac{N_1}{N})^2$
System average travel cost	$\bar{c}^{\,\text{n}}$	$c^*\,(\frac{N_1}{N})^2$	$\sum_\tau g_\tau\, \bar{c}_\tau^{\text{n}}$
Type average queuing delay	$\bar{t}_1^{\text{q}}$	$\frac{c^*}{2\alpha}\,\frac{N_1}{N}$	$\sum_u \mathbb{P}_\tau[u]\, \bar{t}_\tau^{\text{q}}[u]$
	$\bar{t}_1^{\text{q}}$	0	with (B.15)
Type average travel cost	$\bar{c}_1^{\text{n}}$	$c^*\,\frac{N_1}{N}$	$\frac{1}{\bar{u}_\tau}\sum_u \mathbb{P}_\tau[u]\, c_\tau^*[u]$
	$\bar{c}_2^{\text{n}}$	0	with (B.15)

Using this result, we can easily derive the departure window of Type 2 as

$$t_2^{\text{start}} = t^{\text{start}}, \quad t_2^{\text{end}} = \frac{\alpha - \beta}{\alpha}\, t_2^* + \frac{\beta}{\alpha}\, t^{\text{start}}. \tag{B.10}$$

Under NOM, the normalized travel costs of the two types are

$$c_1^* = c^*, \tag{B.11}$$

$$c_2^* = c^* - \beta\,(t_1^* - t_2^*), \tag{B.12}$$

and their average queuing delays are

$$\bar{t}_1^{\text{q}} = \frac{c^*}{2\,\alpha} + \frac{N_2}{N_1}\,\frac{\beta}{2\,\alpha}\,(t_1^* - t_2^*), \tag{B.13}$$

$$\bar{t}_2^{\text{q}} = \frac{c^*}{2\,\alpha} - \frac{\beta}{2\,\alpha}\,(t_1^* - t_2^*). \tag{B.14}$$

Since commuters share the same VOT process, the overall departure pattern with respect to VOT values remains the same as in Section B.1. Therefore, the average queuing delay and travel cost follow the same forms as (B.8) and (B.9). For commuters of type τ, they are

$$\bar{t}_\tau^{\text{q}}[u] = \begin{cases} \bar{t}_\tau^{\text{q}}, & u < \hat{u}, \\ \hat{r}\,\bar{t}_\tau^{\text{q}}, & u = \hat{u}, \\ 0, & u > \hat{u}, \end{cases} \tag{B.15}$$

$$c_\tau^*[u] = \begin{cases} u\,c_\tau^*, & u < \hat{u}, \\ u\,c_\tau^*\left[1 - (1-\hat{r})\,\frac{s}{2\,s^{\text{fast}}}\,\mathbb{P}[\hat{u}]\right], & u = \hat{u}, \\ u\,c_\tau^*\left(1 - \frac{s}{2\,s^{\text{fast}}}\,\mathbb{P}[u]\right), & u > \hat{u}, \end{cases} \tag{B.16}$$

where c^*_τ and $\bar{t}^{\mathrm{q}}_\tau$ take the values in (B.11)–(B.14).

B.2.2. Double peak

When there are two peaks, the two commuter types have two exclusive departure windows. Let t^{bnd} denote the boundary arrival time (the arrival time of the last Type 2 commuter and the first Type 1 commuter). Note that (B.3) can also be written as a function of arrival time $t^{\mathrm{a}} = t + \frac{q(t)}{s}$ [63]. Then, the starting, ending, and boundary arrival times satisfy the following conditions:

$$\begin{cases} \left(t^{\mathrm{end}} - t^{\mathrm{start}}\right) s = N, \quad \left(t^{\mathrm{bnd}} - t^{\mathrm{start}}\right) s = N_2, \quad \left(t^{\mathrm{end}} - t^{\mathrm{bnd}}\right) s = N_1, \\ \beta\left(t^*_2 - t^{\mathrm{start}}\right) - \gamma\left(t^{\mathrm{bnd}} - t^*_2\right) = \gamma\left(t^{\mathrm{end}} - t^*_1\right) - \beta\left(t^*_1 - t^{\mathrm{bnd}}\right). \end{cases} \tag{B.17}$$

The last equation in (B.17) states that the queue length observed by the last Type 2 commuter is the same as that observed by the first Type 1 commuter. Solving the system of equations (B.17) yields

$$\begin{aligned} t^{\mathrm{start}} &= \frac{1}{2}\left(t^*_1 + t^*_2 - \frac{c^*}{\beta} - \frac{N_2}{s}\right), \quad t^{\mathrm{end}} = \frac{1}{2}\left(t^*_1 + t^*_2 + \frac{c^*}{\gamma} + \frac{N_1}{s}\right), \\ t^{\mathrm{bnd}} &= \frac{1}{2}\left(t^*_1 + t^*_2 - \frac{\gamma}{\beta+\gamma}\frac{N_1}{s} + \frac{\beta}{\beta+\gamma}\frac{N_2}{s}\right). \end{aligned}$$

Accordingly, the normalized travel cost of the two groups are

$$c^*_1 = \frac{1}{2}\left[c^* + \gamma\left(t^*_2 - t^*_1 + \frac{N_1}{s}\right)\right], \tag{B.18}$$

$$c^*_2 = \frac{1}{2}\left[c^* + \beta\left(t^*_2 - t^*_1 + \frac{N_2}{s}\right)\right], \tag{B.19}$$

and the system normalized average travel cost is

$$\bar{c}^{\mathrm{n}} = \frac{1}{2}\left[c^* + \frac{\gamma N_1 + \beta N_2}{N}\left(t^*_2 - t^*_1\right) + \frac{\gamma N_1^2 + \beta N_2^2}{N\, s}\right]. \tag{B.20}$$

With some algebra, we derive the average queuing delay for each commuter type as

$$\bar{t}^{\mathrm{q}}_1 = \frac{c^*}{2\,\alpha} + \frac{\gamma}{2\,\alpha}\left(t^*_2 - t^*_1 + \frac{N_1}{s}\right) - \frac{s}{8\,\alpha\,N_1}\left[\beta\left(t^*_2 - t^*_1 + \frac{c^*}{\gamma} - \frac{N_1}{s}\right)^2 + \gamma\left(t^*_2 - t^*_1 + \frac{c^*}{\gamma} + \frac{N_1}{s}\right)^2\right], \tag{B.21}$$

$$\bar{t}^{\mathrm{q}}_2 = \frac{c^*}{2\alpha} + \frac{\beta}{2\alpha}\left(t^*_2 - t^*_1 + \frac{N_2}{s}\right) - \frac{s}{8\alpha N_2}\left[\beta\left(t^*_2 - t^*_1 + \frac{c^*}{\beta} + \frac{N_2}{s}\right)^2 + \gamma\left(t^*_2 - t^*_1 + \frac{c^*}{\beta} - \frac{N_2}{s}\right)^2\right]. \tag{B.22}$$

Finally, the system-level average queuing delay is

$$\begin{aligned} \bar{t}^{\mathrm{q}} = &\frac{c^*}{2\,\alpha} + \frac{\gamma\,N_1^2 + \beta\,N_2^2}{2\,\alpha\,N\,s} - \frac{(\beta+\gamma)\,s}{4\,\alpha\,N}\left(t^*_2 - t^*_1\right)^2 \\ &- \frac{s}{8\,\alpha\,N}\left\{\beta\left[\left(\frac{c^*}{\gamma} - \frac{N_1}{s}\right)^2 + \left(\frac{c^*}{\beta} + \frac{N_2}{s}\right)^2\right] + \gamma\left[\left(\frac{c^*}{\gamma} + \frac{N_1}{s}\right)^2 + \left(\frac{c^*}{\beta} - \frac{N_2}{s}\right)^2\right]\right\}. \end{aligned} \tag{B.23}$$

Similar to the single-peak scenario, the performance measures under `TOLL` are in line with the forms (B.8) and (B.9) but differ in the type-specific values of $\bar{t}^{\mathrm{q}}_\tau$ and c^*_τ. Hence, they share the same formulae of (B.15) and (B.16) though with type-specific travel cost and average queuing delay given in (B.18)–(B.22). The system average queuing delay is (B.23) multiplied by a factor of s^{slow}/s.

B.2.3. Precongestion

In this scenario, the two commuter types have disjoint departure windows. Since the number of Type 2 commuters is below the slow lane capacity, all of them can traverse the bottleneck in the time interval that includes t_2^*. Therefore, for the sake of simplicity, we consider that Type 2 commuters do not endure any queuing or schedule delay cost. Type 1 commuters then start departing and generate bottleneck congestion independently. There, the performance measures can be easily derived as reported in Table B.6.

C. Karma mechanism: theoretical framework

In this section, we revisit the *karma mechanism* firstly introduced in [47] that forms the foundation of the analysis in Section 3. In brief, karma is used to repeatedly allocate a resource to a group of competing agents, viz. commuters in `CARMA`, over an infinite time horizon. At each time step, an agent endowed with an integer quantity $k \in \mathbb{N}$, called *karma*, can submit an integer karma bid $b \in \mathcal{B}[k] = \{0, \dots, k\}$. Apart from the karma, each agent also features an urgency state $u \in \mathcal{U}_\tau = \{u_1, \dots, u_{U_\tau}\}$, $u > 0$, viz. the VOT in `CARMA`, evolving according to an exogenous irreducible Markov chain $\phi_\tau[u^+ \mid u]$ for every agent type $\tau \in \Gamma \subset \mathbb{N}$.

Formally, the karma mechanism is modeled as a *Dynamic Population Game (DPG)* [28]. Namely, the number of agents N is assumed large, thus they can be approximated by a continuum of mass. The distribution of agent types is compactly denoted by $g \in \Delta(\Gamma)$, where $g_\tau \in [0, 1]$ is the mass of agents in type $\tau \in \Gamma$. Accordingly, the time-varying joint *type-state distribution* is given by $d \in \mathcal{D} = \left\{ d \in \mathbb{R}_+^{|\Gamma| \times |\mathcal{X}|} \mid \sum_{u \in \mathcal{U}_\tau} \sum_{k \in \mathbb{N}} d_\tau[u, k] = g_\tau, \text{ for all } \tau \in \Gamma \right\}$, where $d_\tau[u, k]$ denotes the mass of agents in the *static type* τ and *dynamic state* $[u, k] \in \mathcal{X}_\tau$, and $\mathcal{X}_\tau = \mathcal{U}_\tau \times \mathbb{N}$.

At each time step, each agent can choose an action a from a finite state-dependent discrete set $\mathcal{A}[u, k]$. The action includes the agent's bid b as well as other decisions, viz. the departure time in `CARMA`. Agents of the same type τ follow the homogeneous randomized policy $\pi_\tau : \mathcal{X}_\tau \to \Delta(\mathcal{A}[u, k])$, where $\pi_\tau[a \mid u, k]$ denotes the probabilistic weight that these agents place on action a when in state $[u, k]$. The concatenation of the policies of all types $\pi = (\pi_\tau)_{\tau \in \Gamma}$ is simply referred to as the *policy*, and the space of policies is denoted by Π.

In the DPG, the tuple of type-state distribution and policy (d, π) is referred to as the *social state* because it gives a macroscopic description of the distribution of agents in the population as well as their behaviors. Hence, all the outcomes at each time step are functions of (d, π). Let $\kappa[k^+ \mid k, a](d, \pi)$ be the *karma transition function* that describes how the agent's karma changes between two consecutive time steps given its current karma k and action a. Then, together with the urgency transition function $\phi_\tau[u^+ \mid u]$, the joint *state transition function* is given by $\rho_\tau[u^+, k^+ \mid u, k, a](d, \pi) = \phi_\tau[u^+ \mid u]\, \kappa[k^+ \mid k, a](d, \pi)$. Moreover, we define $\zeta_\tau[u, a](d, \pi)$ as the *immediate reward function* of each type τ agent in urgency u taking action a. Both the immediate reward and the karma transition are specified in Section 3. Yet, two conditions are required to ensure that the karma mechanism is well defined.

Assumption 4 (Continuity)**.** The immediate reward function $\zeta_\tau[u, a](d, \pi)$ and the karma transition function $\kappa[k^+ \mid k, a](d, \pi)$ are continuous in the social state (d, π).

Assumption 5 (Karma preservation in expectation)**.** Karma is preserved in expectation for all (d, π), i.e., $\mathbb{E}[k^+] = \mathbb{E}[k]$, which expands to

$$\sum_{\tau \in \Gamma} \sum_{u \in \mathcal{U}_\tau} \sum_{k \in \mathbb{N}} d_\tau[u, k] \sum_{a \in \mathcal{A}[u,k]} \pi_\tau[a \mid u, k] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid k, a](d, \pi)\, k^+ = \sum_{\tau \in \Gamma} \sum_{u \in \mathcal{U}_\tau} \sum_{k \in \mathbb{N}} d_\tau[u, k]\, k. \tag{C.1}$$

Refer to [47] for an in-depth discussion on the assumptions and functions introduced above.

Given the social state, each agent faces a Markov Decision Process (MDP). Specifically, the expected reward of the agents of type τ is given by

$$R_\tau[u, k](d, \pi) = \sum_{a \in \mathcal{A}[u,k]} \pi_\tau[a \mid u, k]\, \zeta_\tau[u, a](d, \pi), \tag{C.2}$$

and the state transition follows

$$P_\tau[u^+, k^+ \mid u, k](d, \pi) = \sum_{a \in \mathcal{A}[u,k]} \pi_\tau[a \mid u, k]\, \rho_\tau[u^+, k^+ \mid u, k, a](d, \pi). \tag{C.3}$$

Accordingly, the expected return in the infinite horizon, also known as the value function, is derived as

$$V_\tau[u, k](d, \pi) = R_\tau[u, k](d, \pi) + \delta \sum_{u^+ \in \mathcal{U}_\tau} \sum_{k^+ \in \mathbb{N}} P_\tau[u^+, k^+ \mid u, k](d, \pi)\, V_\tau[u^+, k^+](d, \pi), \tag{C.4}$$

where $\delta \in (0, 1]$ is the future discount factor. A smaller δ models a more myopic agent behavior. To describe the rational decision of each agent, we also need to define the state-action value function, also called Q-function, as

$$Q_\tau[u,k,a](d,\pi) = \zeta_\tau[u,a](d,\pi) + \delta \sum_{u^+\in\mathcal{U}_\tau} \sum_{k^+\in\mathbb{N}} \rho_\tau[u^+,k^+ \mid u,k,a](d,\pi)\, V_\tau[u^+,k^+](d,\pi). \tag{C.5}$$

Then, to maximize the long-term return, each agent chooses a policy based on the *best response* correspondence, given by

$$B_\tau[u,k](d,\pi) \quad = \quad \left\{ \sigma \in \Delta(\mathcal{A}[u,k]) \,\middle|\, \text{for all } \sigma' \in \Delta(\mathcal{A}[u,k]), \sum_{a\in\mathcal{A}[u,k]} (\sigma[a]-\sigma'[a])\, Q_\tau[u,k,a](d,\pi) \geq 0 \right\}.$$

By definition, $B_\tau[u,k](d,\pi)$ gives the set of randomized individual actions that maximize the Q-function, viz. the long-term reward, in state $[u,k]$ for agents of type τ. The above framework models the rational decision-making process that each agent undergoes to select their strategy.

We are finally ready to formally define the equilibrium concept that we consider.

Definition 4 (Stationary Nash Equilibrium (SNE))**.** A Stationary Nash Equilibrium is a social state $(d^*,\pi^*) \in \mathcal{D}\times\Pi$ such that, for all $[\tau,u,k]\in\Gamma\times\mathcal{U}_\tau\times\mathbb{N}$,

$$d^*_\tau[u,k] = \sum_{u^-\in\mathcal{U}_\tau} \sum_{k^-\in\mathbb{N}} d^*_\tau[u^-,k^-]\; P_\tau[u,k \mid u^-,k^-](d^*,\pi^*), \tag{C.6}$$

$$\pi^*_\tau[\cdot \mid u,k] \in B_\tau[u,k](d^*,\pi^*). \tag{C.7}$$

We conclude this section by stating the conditions for a karma mechanism to guarantee the existence of a SNE (see [47] for the proof).

Theorem 6 (Existence of SNE, Th. 1 [47])**.** *Let Assumption 4 and 5 hold. Then for every $\bar{k}\in\mathbb{N}$, there exists a SNE (d^*,π^*) satisfying $\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d^*_\tau[u,k]\,k = \bar{k}$, where $\bar{k}$ is the average amount of karma per agent in the system.*

D. Proofs

D.1. Proof of Proposition 1

The proof is an application of Theorem 6, thus we must prove that Assumptions 4 and 5 are verified. It is straightforward to verify that, with the continuous approximation ψ^ϵ, all the expressions in the derivation of $\zeta_\tau[u,t,b](d,\pi)$ and $\kappa[k^+ \mid k,t,b](d,\pi)$ are continuous in (d,π). Thus, Assumption 4 is satisfied.

Assumption 5 is satisfied if (C.1) holds for the setup considered. Omitting the dependency on (d,π), we have

$$\begin{aligned}
&\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d_\tau[u,k]\sum_{t\in\mathcal{T}}\sum_{b\in\mathcal{B}[k]}\pi_\tau[t,b\mid u,k]\sum_{k^+\in\mathbb{N}}\kappa[k^+\mid k,t,b]\,k^+\\
&\overset{(9),(10)}{=}\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d_\tau[u,k]\sum_{t\in\mathcal{T}}\sum_{b\in\mathcal{B}[k]}\pi_\tau[t,b\mid u,k]\sum_{o\in\mathcal{O}}\psi^\epsilon[o\mid t,b]\sum_{k^+\in\mathbb{N}}\mathbb{P}[k^+\mid k,t,b,o]\,k^+\\
&\overset{(8)}{=}\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d_\tau[u,k]\sum_{t\in\mathcal{T}}\sum_{b\in\mathcal{B}[k]}\pi_\tau[t,b\mid u,k]\left(\psi^\epsilon[\text{fast}\mid t,b]\,(k-b+\bar{p})+\psi^\epsilon[\text{slow}\mid t,b]\,(k+\bar{p})\right)\\
&=\bar{p}+\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d_\tau[u,k]\,k-\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d_\tau[u,k]\sum_{t\in\mathcal{T}}\sum_{b\in\mathcal{B}[k]}\pi_\tau[t,b\mid u,k]\,\psi^\epsilon[\text{fast}\mid t,b]\,b\\
&\overset{(7)}{=}\sum_{\tau\in\Gamma}\sum_{u\in\mathcal{U}_\tau}\sum_{k\in\mathbb{N}} d_\tau[u,k]\,k.
\end{aligned}$$

The equality from the second to the third line is due to $(1-f)\lfloor\bar{p}\rfloor+f\lceil\bar{p}\rceil=\bar{p}$. We conclude the proof by invoking Theorem 6. □

D.2. Proof of Proposition 2

For ease of presentation, we consider two income types $\tau\in\Gamma=\{\text{low},\text{high}\}$, and without loss of generality let $\lambda_{\text{low}}=1$, $\lambda_{\text{high}}=\lambda>1$. The proof holds for an arbitrary number of types using analogous arguments.

We prove the statement by constructing an equitable `CARMA` SNE. Consider an auxiliary income-homogeneous setting with the same total number of commuters where all are of the low income type, and let $(d^{*,\text{hom}},\pi^{*,\text{hom}})$ be a corresponding `CARMA` SNE, which is guaranteed to exist by Proposition 1. Construct the social state $(d^{*,\text{eq}},\pi^{*,\text{eq}})$ as

$$\begin{aligned}
&d_\tau^{*,\text{eq}}=g_\tau\,d^{*,\text{hom}},\ \text{for all }\tau\in\Gamma, \qquad \pi_{\text{low}}^{*,\text{eq}}=\pi^{*,\text{hom}},\\
&\pi_{\text{high}}^{*,\text{eq}}[a\mid\lambda\,u,k]=\pi^{*,\text{hom}}[a\mid u,k],\ \text{for all }[u,k,a]\in\mathcal{U}_{\text{low}}\times\mathbb{N}\times\mathcal{A}[k].
\end{aligned}$$

The remainder of the proof is devoted to showing that $(d^{*,\text{eq}},\pi^{*,\text{eq}})$ is a SNE in the heterogeneous setting, thus it satisfies (C.6)–(C.7). (C.6) is satisfied since $d^{*,\text{hom}}$ is stationary when the whole population follows $\pi^{*,\text{hom}}$, and thus $d^{*,\text{eq}}$ is also stationary. We therefore focus on showing that (C.7) holds, i.e., that $\pi^{*,\text{eq}}$ is optimal for both commuter types. Observe that the elementary constituents of the commuter MDPs, i.e., the immediate reward function $\zeta_\tau(d,\pi)$ and karma transition function $\kappa_\tau(d,\pi)$, depend on (d,π) only through the bid distribution $\nu(d,\pi)$. It is straight-forward to show that $\nu(d^{*,\text{eq}},\pi^{*,\text{eq}})=\nu(d^{*,\text{hom}},\pi^{*,\text{hom}})$, therefore the low income commuters face an identical strategic problem under $(d^{*,\text{eq}},\pi^{*,\text{eq}})$ and $(d^{*,\text{hom}},\pi^{*,\text{hom}})$ in which $\pi_{\text{low}}^{*,\text{eq}}$ is optimal. For the high income commuters, it holds for all

$[u,k,a,u^+,k^+] \in \mathcal{U}_{\text{low}} \times \mathbb{N} \times \mathcal{A}[k] \times \mathcal{U}_{\text{low}} \times \mathbb{N}$ at $(d^{*,\text{eq}},\pi^{*,\text{eq}})$ (which we omit from the notation)

$$\zeta_{\text{high}}[\lambda\, u, a] = \lambda\, \zeta_{\text{low}}[u,a], \quad R_{\text{high}}[\lambda\, u,k] = \lambda\, R_{\text{low}}[u,k],$$

$$\rho_{\text{high}}[\lambda\, u^+, k^+ \mid \lambda\, u, k, a] = \rho_{\text{low}}[u^+, k^+ \mid u,k,a], \quad P_{\text{high}}[\lambda\, u^+, k^+ \mid \lambda\, u, k] = P_{\text{low}}[u^+,k^+ \mid u,k],$$

$$V_{\text{high}}[\lambda\, u,k] = \lambda\, V_{\text{low}}[u,k], \tag{D.1a}$$

$$Q_{\text{high}}[\lambda\, u,k,a] = \lambda\, Q_{\text{low}}[u,k,a]. \tag{D.1b}$$

Since $\pi^{*,\text{hom}}$ is optimal with respect to $Q_{\text{low}}(d^{*,\text{eq}},\pi^{*,\text{eq}})$, and $Q_{\text{high}}(d^{*,\text{eq}},\pi^{*,\text{eq}})$ differs from $Q_{\text{low}}(d^{*,\text{eq}},\pi^{*,\text{eq}})$ by the constant scaling factor $\lambda > 1$, it follows that $\pi^{*,\text{eq}}_{\text{high}}$ is also optimal. Finally, it holds that

$$\begin{aligned}
\bar{c}^{\text{n}}_{\text{high}}(d^{*,\text{eq}},\pi^{*,\text{eq}}) &= -\frac{1}{g_{\text{high}}\,\bar{u}_{\text{high}}} \sum_{u\in\mathcal{U}_{\text{high}}} \sum_{k\in\mathbb{N}} d^{*,\text{eq}}_{\text{high}}[u,k]\; R_{\text{high}}[u,k](d^{*,\text{eq}},\pi^{*,\text{eq}}) \\
&= -\frac{1}{\lambda\,\bar{u}_{\text{low}}} \sum_{u\in\mathcal{U}_{\text{low}}} \sum_{k\in\mathbb{N}} d^{*,\text{hom}}[u,k]\; \lambda\, R_{\text{low}}[u,k](d^{*,\text{eq}},\pi^{*,\text{eq}}) \\
&= -\frac{1}{g_{\text{low}}\,\bar{u}_{\text{low}}} \sum_{u\in\mathcal{U}_{\text{low}}} \sum_{k\in\mathbb{N}} d^{*,\text{eq}}_{\text{low}}[u,k]\; R_{\text{low}}[u,k](d^{*,\text{eq}},\pi^{*,\text{eq}}) = \bar{c}^{\text{n}}_{\text{low}}(d^{*,\text{eq}},\pi^{*,\text{eq}}),
\end{aligned}$$

which concludes the proof. □

D.3. Proof of Proposition 3

Analogously to the proof of Proposition 2, without loss of generality we consider two income types $\tau \in \Gamma = \{\text{low}, \text{high}\}$, and let $\lambda_{\text{low}} = 1$, $\lambda_{\text{high}} = \lambda > 1$. Mutandi mutandis from [64, Proposition 4.1.2], the following holds for all $[\tau,u,k] \in \Gamma \times \mathcal{U}_\tau \times \mathbb{N}$,

$$\bar{c}^{\text{n}}_\tau[u,k](d^{*,\delta},\pi^{*,\delta}) = -(1-\delta)\,\frac{V_\tau[u,k](d^{*,\delta},\pi^{*,\delta})}{\bar{u}_\tau} + (1-\delta)\,\frac{h_\tau[u,k](d^{*,\delta},\pi^{*,\delta})}{\bar{u}_\tau} + O((1-\delta)^2), \tag{D.2}$$

where $\bar{c}^{\text{n}}_\tau[u,k](d^{*,\delta},\pi^{*,\delta})$ is the normalized long-term average travel cost *starting in state* $[u,k]$, and $h_\tau[u,k](d^{*,\delta},\pi^{*,\delta})$, defined as in [64, Eq. 4.16], is a so-called *bias term* whose order of magnitude does not depend on δ. Thus, we have for all $[u,k] \in \mathcal{U}_{\text{low}} \times \mathbb{N}$

$$\begin{aligned}
&\lim_{\delta\to 1} \left[\bar{c}^{\text{n}}_{\text{high}}[\lambda\, u,k](d^{*,\delta},\pi^{*,\delta}) - \bar{c}^{\text{n}}_{\text{low}}[u,k]\Big(d^{*,\delta},\pi^{*,\delta})\right] \\
&= \lim_{\delta\to 1}\left[(1-\delta)\left(\frac{h_{\text{high}}[\lambda\, u,k]}{\bar{u}_{\text{high}}} - \frac{h_{\text{low}}[u,k]}{\bar{u}_{\text{low}}}\right) + O((1-\delta)^2)\right] = 0,
\end{aligned}$$

where we used (D.1a) and the optimality of $\pi^{*,\delta}_{\text{high}}$ and $\pi^{*,\delta}_{\text{low}}$. This proves the result conditioned that commuters of both income groups start from the same (normalized) state $[u,k] \in \mathcal{U}_{\text{low}} \times \mathbb{N}$.

To conclude the proof we establish that for all $[\tau,u,k,u',k'] \in \Gamma \times \mathcal{U}_\tau \times \mathbb{N} \times \mathcal{U}_\tau \times \mathbb{N}$ it holds that $\lim_{\delta\to 1}\left[\bar{c}^{\text{n}}_\tau[u,k](d^{*,\delta},\pi^{*,\delta}) - \bar{c}^{\text{n}}_\tau[u',k'](d^{*,\delta},\pi^{*,\delta})\right] = 0$, i.e., $\bar{c}^{\text{n}}_\tau[u,k](d^{*,\delta},\pi^{*,\delta})$ does not depend on $[u,k]$ as $\delta \to 1$. Using (D.2) we have

$$\lim_{\delta\to 1}\left[\bar{c}^{\text{n}}_\tau[u,k](d^{*,\delta},\pi^{*,\delta}) - \bar{c}^{\text{n}}_\tau[u',k'](d^{*,\delta},\pi^{*,\delta})\right] = \lim_{\delta\to 1} \frac{(1-\delta)}{\bar{u}_\tau}\left(V_\tau[u',k'](d^{*,\delta},\pi^{*,\delta}) - V_\tau[u,k](d^{*,\delta},\pi^{*,\delta})\right).$$

Then, from Lemma 7, it follows that for all $[\tau, u, k, k'] \in \Gamma \times \mathcal{U}_\tau \times \mathbb{N} \times \mathbb{N}$ it holds

$$\lim_{\delta \to 1} \left[\bar{c}^{\mathrm{n}}_\tau[u,k](d^{*,\delta}, \pi^{*,\delta}) - \bar{c}^{\mathrm{n}}_\tau[u,k'](d^{*,\delta}, \pi^{*,\delta}) \right] > 0 \Rightarrow k' > k. \tag{D.3}$$

Moreover, it is well known that the long-term average cost in a Markov chain is identical for all initial states belonging to the same communicating class, i.e., it differs per initial state only if the Markov chain has multiple communicating classes, see, e.g., [64, Chapter 4.2]. Since the VOT Markov chain ϕ_τ is irreducible this is only possible if there are more than one communicating class in the karma states. Suppose that this is the case, i.e., there exist $\tau \in \Gamma$, $u' \in \mathcal{U}_\tau$, $k', k'' \in \mathbb{N}$ such that $\lim_{\delta \to 1} \left[\bar{c}^{\mathrm{n}}_\tau[u',k'](d^{*,\delta}, \pi^{*,\delta}) - \bar{c}^{\mathrm{n}}_\tau[u',k''](d^{*,\delta}, \pi^{*,\delta}) \right] > 0$. Due to (D.3) we must have that $k'' > k'$, and similarly for all $k \geq k'' > k'$ we must have $\lim_{\delta \to 1} \left[\bar{c}^{\mathrm{n}}_\tau[u',k'](d^{*,\delta}, \pi^{*,\delta}) - \bar{c}^{\mathrm{n}}_\tau[u',k](d^{*,\delta}, \pi^{*,\delta}) \right] > 0$. Thus, it must hold that k' does not communicate with any $k \geq k''$. Consider a candidate policy $\tilde{\pi}_\tau$ in which commuters bid 0 at all karma states $k \in [k', k'')$, and at karma states $k \geq k''$ follow policy $\pi^{*,\delta}_\tau$ that yields $\lim_{\delta \to 1} \bar{c}^{\mathrm{n}}_\tau[u',k](d^{*,\delta}, \pi^{*,\delta})$. By Lemma 8 this policy guarantees that some $k \geq k''$ is reachable from k', after which a lower long-term average cost can be obtained. Therefore, for $\delta \to 1$ this benefit will outweigh any transient costs incurred while reaching $k \geq k''$ from k', hence $\tilde{\pi}_\tau$ will achieve a higher $V_\tau[u', k']$ than $\pi^{*,\delta}_\tau$ that obtains $\lim_{\delta \to} \bar{c}^{\mathrm{n}}_\tau[u',k'](d^{*,\delta}, \pi^{*,\delta})$. This leads to a contradiction since $\pi^{*,\delta}_\tau$ is an equilibrium policy. Thus, there is only one communicating class in the karma states and this concludes the proof. □

D.4. Supporting lemmas for Proposition 4

Lemma 7. *At a CARMA SNE (d^*, π^*), the value function is non-decreasing in karma, i.e., for all $[\tau, u, k] \in \Gamma \times \mathcal{U} \times \mathbb{N}$,*

$$V_\tau[u, k+1](d^*, \pi^*) \geq V_\tau[u, k](d^*, \pi^*). \tag{D.4}$$

Proof. By (C.7) $V_\tau(d^*, \pi^*)$ is an optimal value function, therefore it is the unique fixed point of the contraction mapping given for all $[\tau, u, k] \in \Gamma \times \mathcal{U}_\tau \times \mathbb{N}$ by, see, e.g., [64, Chapter 1.4],

$$\begin{aligned} V^{(l+1)}_\tau[u,k] &= \max_{t \in \mathcal{T},\, b \leq k} \left\{ \zeta_\tau[u,t,b] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid k, t, b]\, V^{(l)}_\tau[u^+, k^+] \right\} \\ &= \max_{t \in \mathcal{T},\, b \leq k} \left\{ \zeta_\tau[u,t,b] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u] \sum_{\Delta k \in \mathbb{N}} \mathbb{P}[\Delta k \mid t, b]\, V^{(l)}_\tau[u^+, k + \Delta k] \right\}, \end{aligned} \tag{D.5}$$

where we used $\mathbb{P}[\Delta k = k^+ - k \mid t, b] = \mathbb{P}[\Delta k = k^+ - k \mid k, t, b] := \kappa[k^+ \mid k, t, b]$ and omitted (d^*, π^*) from the notation. Equation (D.5) is the well-known Bellman iteration.

We then proceed by induction over l. Initialize with an arbitrary non-decreasing value function in karma, e.g., $V_\tau^{(0)}[u,k]=0$ for all $[\tau,u,k]\in\mathcal{T}\times\mathcal{U}_\tau\times\mathbb{N}$. As induction hypothesis assume that $V_\tau^{(l)}[u,k+1]\geq V_\tau^{(l)}[u,k]$ for all $[\tau,u,k]\in\mathcal{T}\times\mathcal{U}_\tau\times\mathbb{N}$. Then it holds for $l+1$,

$$\begin{aligned}
V_\tau^{(l+1)}[u,k+1] &= \max_{t\in\mathcal{T},\, b\leq k+1}\left\{\zeta_\tau[u,t,b]+\delta\sum_{u^+\in\mathcal{U}_\tau}\phi_\tau[u^+\mid u]\sum_{\Delta k\in\mathbb{N}}\mathbb{P}[\Delta k\mid t,b]\,V_\tau^{(l)}[u^+,k+1+\Delta k]\right\}\\
&\geq \max_{t\in\mathcal{T},\, b\leq k}\left\{\zeta_\tau[u,t,b]+\delta\sum_{u^+\in\mathcal{U}_\tau}\phi_\tau[u^+\mid u]\sum_{\Delta k\in\mathbb{N}}\mathbb{P}[\Delta k\mid t,b]\,V_\tau^{(l)}[u^+,k+1+\Delta k]\right\}\\
&\geq \max_{t\in\mathcal{T},\, b\leq k}\left\{\zeta_\tau[u,t,b]+\delta\sum_{u^+\in\mathcal{U}_\tau}\phi_\tau[u^+\mid u]\sum_{\Delta k\in\mathbb{N}}\mathbb{P}[\Delta k\mid t,b]\,V_\tau^{(l)}[u^+,k+\Delta k]\right\}=V_\tau^{(l+1)}[u,k]
\end{aligned}$$

as needed. Letting $l\to\infty$ establishes the result. □

Lemma 8. *At a* CARMA *SNE* (d^*,π^*)*, the total karma redistributed to commuters is strictly positive. Due to the uniform redistribution scheme, this is equivalent to a positive average payment* $\bar{p}(d^*,\pi^*)>0$.

Proof. We omit (d^*,π^*) for notational convenience. Suppose, for the sake of contradiction, that $\bar{p}=0$. Then it must hold that $\pi^*_\tau[t,b>0\mid u,k]=0$ for all $[\tau,u,k,t]\in\Gamma\times\mathcal{U}_\tau\times\mathbb{N}\times\mathcal{T}$ with $d^*_\tau[u,k]>0$. Moreover, since with $\bar{p}=0$ there are no karma dynamics and the VOT dynamics are exogenous, using a similar argument as in the proof of Proposition 4 it is straightforward to show that the value function $V_\tau[u,k]$ depends on u only and maximizing the state-action value function $Q_\tau[u,k,a]$ is equivalent to maximizing the immediate rewards $\zeta_\tau[u,a]$. It must also hold that $q[t^*]>0$; otherwise action $[t^*,b=0]$ incurs zero immediate cost and strongly dominates all other actions leading to $q[t^*]>0$.

Since all commuters are initially endowed with a positive amount of karma, there must exist $[\tilde{\tau},\tilde{u},\tilde{k}\geq 1]$ for which $d^*_{\tilde{\tau}}[\tilde{u},\tilde{k}]>0$. But action $[t^*,b=1]$ guarantees zero immediate cost and strongly dominates all actions with $b=0$, contradicting that $\pi^*_{\tilde{\tau}}$ is an equilibrium policy and concluding the proof. □

Lemma 9. *Under Assumptions 2 and 3, there is no queue at* $t=t^{\text{start}}$ *at any* CARMA *SNE* (d^*,π^*)*. As a consequence, the expected immediate rewards satisfy for all* $[\tau,u,k]\in\Gamma\times\mathcal{U}_\tau\times\mathbb{N}$,

$$R_\tau[u,k](d^*,\pi^*)\geq -u\,c^*. \tag{D.6}$$

Proof. We omit (d^*,π^*) for notational convenience and first show that there is no queue at $t=t^{\text{start}}$. Suppose, for the sake of contradiction, that $q[t^{\text{start}}]>0$. Then the following conditions must hold:

1. There exists $[\tilde{\tau}, \tilde{u}, \tilde{k}, \tilde{b}] \in \Gamma \times \mathcal{U}_{\tilde{\tau}} \times \mathbb{N} \times \mathcal{B}[\tilde{k}]$ such that $\pi^*_{\tilde{\tau}}[t^{\text{start}}, \tilde{b} \mid \tilde{u}, \tilde{k}] > 0$ and

$$\begin{aligned} Q_{\tilde{\tau}}[\tilde{u}, \tilde{k}, t^{\text{start}}, \tilde{b}] &= -\zeta_{\tilde{\tau}}[\tilde{u}, t^{\text{start}}, \tilde{b}] + \delta \sum_{u^+ \in \mathcal{U}_{\tilde{\tau}}} \phi_{\tilde{\tau}}[u^+ \mid \tilde{u}] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid \tilde{k}, t^{\text{start}}, \tilde{b}]\, V_{\tilde{\tau}}[u^+, k^+] \\ &< -\tilde{u}\, c^* + \delta \sum_{u^+ \in \mathcal{U}_{\tilde{\tau}}} \phi_{\tilde{\tau}}[u^+ \mid \tilde{u}] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid \tilde{k}, t^{\text{start}}, \tilde{b}]\, V_{\tilde{\tau}}[u^+, k^+] \qquad \text{(D.7a)} \\ &\leq -\tilde{u}\, c^* + \delta \sum_{u^+ \in \mathcal{U}_{\tilde{\tau}}} \phi_{\tilde{\tau}}[u^+ \mid \tilde{u}]\, V_{\tilde{\tau}}[u^+, \tilde{k} + \bar{p}], \qquad \text{(D.7b)} \end{aligned}$$

   where we defined $V_{\tilde{\tau}}[u^+, \tilde{k} + \bar{p}] := f\, V_{\tilde{\tau}}[u^+, \tilde{k} + \lceil \bar{p} \rceil] + (1 - f)\, V_{\tilde{\tau}}[u^+, \tilde{k} + \lfloor \bar{p} \rfloor]$. Inequality (D.7a) holds because $c^* = \beta\, (t^* - t^{\text{start}})$ is the queue-free normalized cost at t^{start} and therefore a strictly greater immediate cost is incurred when $q[t^{\text{start}}] > 0$, and (D.7b) holds due to Lemma 7 and that $\tilde{k} + \bar{p}$ is the maximum possible next karma.

2. For all $t \in \{t^{\text{start}} + \Delta, \ldots, t^{\text{end}}\}$, $q[t] > 0$. Otherwise, if $q[t'] = 0$ for some $t' \in \{t^{\text{start}} + \Delta, \ldots, t^{\text{end}}\}$, it is straight-forward to verify that $Q_{\tilde{\tau}}[\tilde{u}, \tilde{k}, t', b = 0]$ is greater or equal to (D.7b), contradicting that $[t^{\text{start}}, \tilde{b}]$ is an equilibrium action.

3. Similarly, for all $t \in \{t^{\text{start}}, \ldots, t^{\text{end}}\}$, the fast lane is fully occupied, i.e., $N^{\text{fast}}[t] = s^{\text{fast}}\, \Delta$, where $N^{\text{fast}}[t]$ is the number of commuters entering the fast lane at time t (with minor loss of generality it is assumed that the parameter ϵ has a negligible effect in under-allocating the fast lane. Formally, $N^{\text{fast}}[t] = s^{\text{fast}}\Delta - O(\epsilon)$, which does not affect the result for $\epsilon \to 0$). Otherwise, if it is not fully occupied for some $t' \in \{t^{\text{start}}, \ldots, t^{\text{end}}\}$, it holds that $b^*[t'] = 0$ and action $[t', b = 0]$ guarantees entry into the fast lane, achieving $Q_{\tilde{\tau}}[\tilde{u}, \tilde{k}, t', b = 0]$ greater or equal to (D.7b) and contradicting that $[t^{\text{start}}, \tilde{b}]$ is an equilibrium action.

4. Let $\hat{t} = \max\{t < t^{\text{start}} \mid q[t] = 0\}$ be the latest queue-free departure time preceding t^{start}. For all $t \in \{\hat{t} + \Delta, \ldots, t^{\text{end}}\}$, since $q[t] = \max\{0, q[t - \Delta] + n^{\text{slow}}[t] - s^{\text{slow}}\, \Delta\} > 0$, we have $N^{\text{slow}}[t] = s^{\text{slow}}\, \Delta + q[t] - q[t - \Delta]$, where $N^{\text{slow}}[t]$ is the number of commuters entering the slow lane at time t.

It follows that the total number of departures in $t \in \{\hat{t} + \Delta, \ldots, t^{\text{end}}\}$ is

$$\begin{aligned} \sum_{t=\hat{t}+\Delta}^{t^{\text{end}}} \left(N^{\text{slow}}[t] + N^{\text{fast}}[t]\right) &\geq \sum_{t=\hat{t}+\Delta}^{t^{\text{end}}} N^{\text{slow}}[t] + \sum_{t=t^{\text{start}}}^{t^{\text{end}}} N^{\text{fast}}[t] \\ &= \sum_{t=\hat{t}+\Delta}^{t^{\text{end}}} \left(s^{\text{slow}}\, \Delta + q[t] - q[t - \Delta]\right) + (t^{\text{end}} - t^{\text{start}} + \Delta)\, s^{\text{fast}} \\ &= (t^{\text{end}} - \hat{t})\, s^{\text{slow}} + q[t^{\text{end}}] + (t^{\text{end}} - t^{\text{start}} + \Delta)\, s^{\text{fast}} \\ &\geq (t^{\text{end}} - t^{\text{start}} + \Delta)\, (s^{\text{slow}} + s^{\text{fast}}) + q[t^{\text{end}}] = N + s\, \Delta + q[t^{\text{end}}] > N, \end{aligned}$$

leading to a contradiction. Thus it must hold that $q[t^{\text{start}}] = 0$.

Next we show (D.6). Suppose, for the sake of contradiction, that there exists some $[\tilde{\tau}, \tilde{u}, \tilde{k}]$ for which $-R_{\tilde{\tau}}[\tilde{u}, \tilde{k}] > \tilde{u}\, c^*$. Then there must exist action $\tilde{a}$ with $\pi^*_{\tilde{\tau}}[\tilde{a} \mid \tilde{u}, \tilde{k}] > 0$ for which $\zeta_{\tilde{\tau}}[\tilde{u}, \tilde{a}] < -\tilde{u}\, c^*$, and

$$Q_{\tilde{\tau}}[\tilde{u}, \tilde{k}, \tilde{a}] < -\tilde{u}\, c^* + \delta \sum_{u^+ \in \mathcal{U}_{\tilde{\tau}}} \phi_{\tilde{\tau}}[u^+ \mid \tilde{u}]\, V_{\tilde{\tau}}[u^+, \tilde{k} + \bar{p}], \tag{D.9}$$

following similar arguments as in (D.7). But since $q[t^{\text{start}}] = 0$, action $[t^{\text{start}}, b = 0]$ achieves $Q_{\tilde{\tau}}[\tilde{u}, \tilde{k}, t^{\text{start}}, b = 0]$ equal to the right-hand side of (D.9). This contradicts that $\tilde{a}$ is an equilibrium action and concludes the proof. □

D.5. Proof of Proposition 4

Lemma 9 directly implies that a weak version of Proposition 4 holds, i.e., $\bar{c}^{\text{n}}_\tau(d^*, \pi^*) \le c^*$ for all $\tau \in \Gamma$. We complete the proof by showing that this inequality can be made strict. We omit (d^*, π^*) for notational convenience. For an arbitrary $\tau \in \Gamma$ suppose, for the sake of contradiction, that $\bar{c}^{\text{n}}_\tau = c^*$. Then, due to Lemma 9 it must hold that $R_\tau[u, k] = -u\, c^*$ for all $[u, k] \in \mathcal{U}_\tau \times \mathbb{N}$ satisfying $d^*_\tau[u, k] > 0$, i.e., $R_\tau[u, k]$ depends on u only and not on k, which we denote by $R_\tau[u, k] = R_\tau[u]$. We show that $V_\tau[u, k] = V_\tau[u]$ must also depend on u only and not on k. It is well known that $V_\tau[u, k]$ is the unique fixed point of the contraction mapping given for all $[u, k] \in \mathcal{U}_\tau \times \mathbb{N}$ by, see, e.g., [64, Chapter 1.4],

$$V^{(l+1)}_\tau[u, k] = R_\tau[u] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u] \sum_{a \in \mathcal{A}[k]} \pi_\tau[a \mid u, k] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid k, a]\, V^{(l)}_\tau[u^+, k^+]. \tag{D.10}$$

Then, we show that $V_\tau[u, k] = V_\tau[u]$ by induction over $l \in \mathbb{N}$. Initialize with any $V^{(0)}_\tau[u, k] = V^{(0)}_\tau[k]$, e.g., $V^{(0)}_\tau[u, k] = 0$ for all $[u, k] \in \mathcal{U}_\tau \times \mathbb{N}$. As induction hypothesis assume that $V^{(l)}_\tau[u, k] = V^{(l)}_\tau[u]$. Then it holds for $l + 1$,

$$\begin{aligned} V^{(l+1)}_\tau[u, k] &= R_\tau[u] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u] \sum_{a \in \mathcal{A}[k]} \pi_\tau[a \mid u, k] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid k, a]\, V^{(l)}_\tau[u^+] \\ &= R_\tau[u] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u]\, V^{(l)}_\tau[u^+] = V^{(l+1)}_\tau[u], \end{aligned}$$

as needed. Similarly, for the state-action value function we have

$$Q_\tau[u, k, a] = \zeta_\tau[u, a] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u] \sum_{k^+ \in \mathbb{N}} \kappa[k^+ \mid k, a]\, V_\tau[u^+] = \zeta_\tau[u, a] + \delta \sum_{u^+ \in \mathcal{U}_\tau} \phi_\tau[u^+ \mid u]\, V_\tau[u^+].$$

Notice that, since the VOT process $\phi_\tau[u^+ \mid u]$ is exogeneous, maximizing $Q_\tau[u, k, a]$ with respect to the action a is equivalent to maximizing the immediate rewards $\zeta_\tau[u, a]$.

Furthermore, it must hold that the type τ commuters never enter the fast lane during $t \in \mathcal{T}^{\text{ben}} := \{t^{\text{start}} + \Delta, \ldots, t^{\text{end}} - \Delta\}$; otherwise they would experience an immediate reward that is strictly greater than $-uc^*$, since c^* equals the normalized queue-free cost at t^{start} or t^{end}.

Simultaneously, due to Lemma 8 a type τ commuter has a non-zero probability of gaining a positive amount of karma every day. Therefore, for any $\tilde{t} \in \mathcal{T}^{\text{ben}}$, there is a non-zero probability of eventually reaching a state $[\tilde{u}, \tilde{k}]$, where $\tilde{k} \geq b^*[\tilde{t}] + 1$, i.e., it holds that $d^*_\tau[\tilde{u}, \tilde{k}] > 0$. At such a state, an action $[\tilde{t}, b^*[\tilde{t}] + 1]$ guarantees entry into the fast lane, and an immediate reward that is strictly greater than $-\tilde{u}\, c^*$, which leads to a contradiction and concludes the proof. □

D.6. Proof of Proposition 5

Given $n[t](d,\pi)$ for all $t \in \mathcal{T}$ and $\bar{p}(d,\pi)$, (18)–(19) constitute a linear system of T equations in the T unknowns $r[t](d,\pi)$ for all $t \in \mathcal{T}$. In vector form, this can be written as $A_{\text{re}}(d,\pi)\, r(d,\pi) = b_{\text{re}}(d,\pi)$, where, omitting (d,π) from the notation,

$$\underline{m} = \left[m[t^1], \cdots, m[t^{\text{min-re}} - \Delta]\right]^\top, \qquad \overline{m} = \left[m[t^{\text{min-re}} + \Delta], \cdots, m[t^T]\right]^\top$$

$$A_{\text{re}} = \begin{bmatrix} I_{\text{min-re}-1} & -\underline{m} & 0_{(\text{min-re}-1)\times(T-\text{min-re})} \\ 0_{(T-\text{min-re})\times(\text{min-re}-1)} & -\overline{m} & I_{T-\text{min-re}} \\ n[1] \;\cdots\; n[t^{\text{min-re}} - \Delta] & n[t^{\text{min-re}}] & n[t^{\text{min-re}} + \Delta] \;\cdots\; n[T] \end{bmatrix}, \quad b_{\text{re}} = \begin{bmatrix} 0_{(T-1)\times 1} \\ \bar{p} \end{bmatrix}.$$

The first $T-1$ rows of $A_{\text{re}}(d,\pi)$ are linearly independent. To verify the linear independence of the last row T, we proceed with Gaussian elimination. In the reduced matrix, row T has all zeros except in the $t^{\text{min-re}}$-th position, which equals $\sum_{t\in\mathcal{T}} m[t]\, n[t](d,\pi) > 0$, where we used $m[t^{\text{min-re}}] = 1$. The inequality holds because $m[t] > 0$ and $n[t](d,\pi) \geq 0$ for all $t \in \mathcal{T}$, and for each $(d,\pi) \in \mathcal{D} \times \Pi$ there exists at least one $\tilde{t} \in \mathcal{T}$ for which $n[\tilde{t}](d,\pi) > 0$ (since $\sum_{t\in\mathcal{T}} n[t](d,\pi) = 1$). This implies that $A_{\text{re}}(d,\pi)$ is invertible. Accordingly, the redistribution vector is unique and reads as $r(d,\pi) = A^{-1}_{\text{re}}(d,\pi)\, b_{\text{re}}(d,\pi)$. It is also continuous in (d,π) as $A^{-1}_{\text{re}}(d,\pi)$ and $b_{\text{re}}(d,\pi)$ are continuous in (d,π). □