Online Appendix

C Calibration Strategy and Numerical Methods

C.1 Law of Motion for Infections

In this appendix, we describe how we use information about parameter values in the canonical SIR model to deduce parameter values for the law of motion (1).

C.1.1 Canonical SIR Model

The canonical SIR model due to Kermack and McKendrick (1927) specifies laws of motion for the population shares of three groups: the "susceptible," the "infected" or "infectives," and the "removed." Their respective population shares at time $t \geq 0$ are denoted by $x(t)$, $\iota(t)$, and $z(t)$, respectively, where $x(t)+\iota(t)+z(t) = 1$.[60] We normalize the mass of the total population at time $t = 0$ to unity.

At time $t = 0$ the population consists of $x(0)$ susceptible persons and a few infected persons, $\iota(0)$. There are no removed persons at this time, $z(0) = 0$. In each instant after time $t = 0$, infected persons transmit the disease to members of the susceptible group and a share of the infected either dies or recovers and develops immunity. Formally,

$$\dot{x}(t) = -b(t)x(t)\iota(t), \quad (12)$$

$$\dot{\iota}(t) = -\dot{x}(t) - (c^d + c^r)\iota(t), \quad (13)$$

$$\dot{z}(t) = (c^d + c^r)\iota(t). \quad (14)$$

Here, $b(t)$ denotes a possibly time-varying infection rate. The extent to which susceptible persons are infected depends on their number, $x(t)$; the infection rate, $b(t)$; and the population share of infected persons. The number of infected persons increases one-to-one with the susceptible persons that get infected, while a share $c \equiv c^d + c^r$ of the infected population dies or recovers; the coefficients c^d and c^r parameterize the flow into death and recovery, respectively.

Consider the case where $b(t)$ is constant at value b. Inspection of equations (12) and (13) reveals that for $bx(0) > c$ the share of infected persons increases until it reaches a maximum when $x(t) = c/b$; thereafter, the share declines. Intuitively, when $x(0)$ falls short of c/b (the "herd immunity level") then there are fewer

[60] We follow the notation introduced by Kermack and McKendrick (1927) except for denoting the share of infected by ι rather than y.

new infections of susceptible persons than outflows from the infected pool due to recoveries and death. As is well known (e.g., Theorem 2.1 in Hethcote, 2000), $x(\infty)$ falls short of the herd immunity level unless $x(0) = c/b = x(\infty)$ and $\iota(0) = 0$.[61]

In the SIR-S model a share γ of the removed population loses immunity and moves back to the susceptible pool. Accordingly, the dynamic system is given by

$$\begin{aligned} \dot{x}(t) &= -b(t)x(t)\iota(t) + \gamma z(t), \\ \dot{\iota}(t) &= b(t)x(t)\iota(t) - cy(t), \\ \dot{z}(t) &= c\iota(t) - \gamma z(t). \end{aligned}$$

In steady state this reduces to

$$\gamma z = bx\iota = c\iota.$$

Calibration. We measure time in days and use information about the spread of COVID-19 in the United States to calibrate the model. We associate $t = 0$ with 16 March 2020, the date at which the median U.S. state closed schools and public health authorities considered further restrictions.[62]

Following Atkeson (2020) and the sources cited therein we assume that the flow rate from the infected to the removed population equals $c = 1/18$, corresponding to an exponentially distributed infection duration that averages 18 days.[63] To calibrate b we rely on information in Ferguson et al. (2020) who argue that the "basic reproduction number" $\mathcal{R}_0 = b/c$ for COVID-19 equals approximately 2.4. This implies $b = 0.1333$.

By March 16, 2020 the U.S. reported a total of 91 COVID-19 deaths out of a population of 328 million.[64] With an infection fatality rate of 0.58 percent (Menachemi et al., 2020) this implies $z(0) = 91/(0.58\% \cdot 328 \cdot 10^6) = 0.4783 \cdot 10^{-4}$.[65] We use the following well-known result (e.g., Theorem 2.1 in Hethcote, 2000):

[61] Note also, from equation (13), that at the beginning of an epidemic with $x(t) \approx 1$ and $z(t) \approx 0$, b approximately equals the growth rate of the number of persons who are or were infected, $\frac{\dot{\iota}(t)+\dot{z}(t)}{\iota(t)+z(t)} = b\frac{x(t)\iota(t)}{\iota(t)+z(t)} \approx b$.

[62] The median state closed restaurants and imposed restrictions on public gatherings on 17 March 2020. Puerto Rico imposed a stay-at-home order on 15 March 2020. See `https://web.csg.org/covid19/executive-orders/`.

[63] Note that $\int_0^\infty ce^{-ct}t\ dt = 1/c$.

[64] See `https://github.com/nytimes/covid-19-data/blob/master/us.csv`.

[65] See Farboodi et al. (2021) for similar calculations.

Proposition 4. Let $s < 0$ denote a start date. In the canonical SIR model with $\iota(s) > 0$ and $b(t) = b > c/x(s)$,

$$1 - z(0) = x(0) + \iota(0) = x(s) + \iota(s) + \frac{c}{b} \ln\left(\frac{x(0)}{x(s)}\right).$$

The long-run share of the susceptible population, $x(\infty)$, solves the equation

$$x(\infty) = x(0) + \iota(0) + \frac{c}{b} \ln\left(\frac{x(\infty)}{x(0)}\right).$$

Since the number of infected or removed persons was negligible before mid March the first set of equalities implies

$$1 - z(0) = 1 + \frac{1}{2.4} \ln\left(\frac{x(0)}{1}\right) \quad \Rightarrow \quad x(0) = 1 - 0.1148 \cdot 10^{-3}, \quad \iota(0) = 0.6696 \cdot 10^{-4}.$$

The second condition implies $x(\infty) = 0.1214$.

C.1.2 Generalized Logistic Model

Given $y(0) = \iota(0) + z(0) = 0.6696 \cdot 10^{-4} + 0.4783 \cdot 10^{-4} = 0.1148 \cdot 10^{-3}$ and $\bar{y} = 1 - x(\infty) = 0.8786$, we choose β and ω in equation (1) (subject to $g(a) = 1$) to best fit the path for $f(t)$ according to the law of motion (1) to the path for $\iota(t)$ in the calibrated SIR model (allowing for an arbitrary factor of proportionality). This yields $\beta = 0.8346 \cdot 10^{-1}$ and $\omega = 0.6662$. Figure 6 illustrates the very close parallels between the predictions of the canonical SIR model (in blue) and the generalized logistic model (in black).

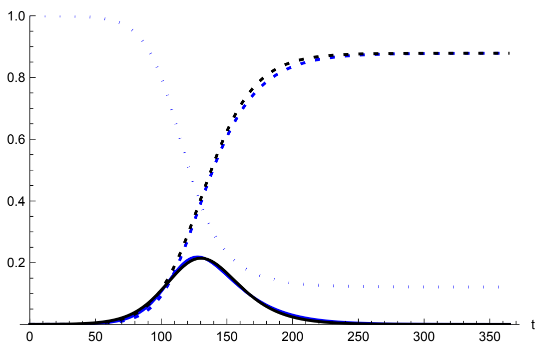

Figure 6: Dynamics in the canonical SIR model (blue) and in the generalized logistic model (black). SIR model: $x(t)$ (dotted), $\iota(t)$ (solid), and $z(t)$ (dashed). Generalized logistic model: $f(t)$ (solid, scaled), and $y(t)$ (dashed).

C.2 Costs of Infection

In this appendix, we discuss the calibration of the parameters representing private and social costs of infection. Recall that $\xi = \zeta \frac{a_i}{a} + (1-\zeta)$. We calibrate ζ based on U.S. estimates of hospitalization costs and the value of life by Bartsch et al. (2020) and Hall et al. (2020), respectively. Bartsch et al. (2020) estimate direct medical costs including follow up expenses (over a year) of \$1.25 trillion under the assumption that eighty percent of the U.S. population are infected. We scale this number reflecting our modified specification of $\bar{y} = 1 - x(\infty)$ (0.8786 rather than 0.8000); this yields (unconditional) costs per capita of about \$4,185. Hall et al. (2020) assess the value of life at \$270,000 per year. With an average remaining life expectancy of 14.5 years every life lost to COVID-19 thus costs \$3,915,000. Conditional on the infection fatality rate of 0.58 percent (Menachemi et al., 2020) and $\bar{y} = 0.8786$ this implies (unconditional) costs due to COVID-19 deaths of \$19,950 per capita. Under the assumption that individuals fully internalize mortality risk but not marginal social medical costs we conclude that $\zeta = 19,950/(19,950 + 4,185) = 0.8266$.

To calibrate ψ based on the dollar amount \$19,950 + \$4,185 we follow the approach of Hall et al. (2020) who quantify willingness to pay for reduced mortality risk. Under our baseline assumptions of an infection risk of 0.8786, an infection fatality rate of 0.58 percent, and relative risk aversion of one we find that an individual would sacrifice a share $1 - \phi = 0.3581$ of consumption to eliminate COVID-19 related mortality risk (neglecting other costs).[66] In the model the annual utility costs of sacrificing this share then equal[67]

$$365 \cdot \{(1 + \ln(a^\star) - a^\star) - (1 + \ln(a^\star \phi) - a^\star)\} = -365 \ln(\phi).$$

[66] Let ℓ denote expected costs due to loss of life (which we calibrate to \$19,950, see above) relative to average consumption (which Hall et al. (2020) set to \$45,000). This ratio, which represents the sacrifice ratio of an agent with linear preferences, equals 0.4433. Hall et al. (2020) derive ϕ from ℓ based on the relationship $\phi = [1 + (r-1)\ell]^{1/(1-r)}$ where r denotes the coefficient of relative risk aversion. For $r = 0$ this yields $1 - \phi = \ell$; for $r \to 1$ (logarithmic utility), $1 - \phi \to 1 - \exp(-\ell)$ or roughly 0.3581. For $r = 2$ we find $1 - \phi = 0.3071$. Hall et al. (2020) assume CRRA preferences, a coefficient of relative risk aversion of two, and mortality risk of 0.44 percent per year. They report that the share of consumption that an individual would sacrifice to eliminate COVID-19 related mortality risk equals 28 percent.

[67] We neglect time discounting as do Hall et al. (2020). Note that only the benefit of economic activity ("consumption") not the cost associated with it ("labor supply") is reduced by the fraction $1 - \phi$.

Since without reductions in activity almost all infections occur during the first year ($\int_0^{365} f(t)dt \approx \bar{y}$) we conclude that the social cost parameter reflecting mortality risk, $\hat{\psi}$, equals $\hat{\psi} = -365 \ln(\phi)/\bar{y} = 184.2$. Adding medical costs we arrive at an estimate for $\psi = \hat{\psi}/\zeta$ of 222.8.[68]

According to the model, the unconditional per-capita utility cost of not adjusting behavior (i.e. $a = a^\star$) is approximately 222.8, and corresponds to a dollar amount of \$19,950 + \$4,185. This allows us to transform welfare measures in utility units into their equivalent dollar amounts.

Endogenous Costs. When ψ is a function of the state we need to modify the calibration. To capture congestion effects we replace ψ in (3) by $\psi^f y(t)(1-(\frac{y(t)}{\bar{y}})^\omega)$ where $\psi^f > 0$. Using

$$\int_0^\infty f(t)y(t)\left(1-\left(\frac{y(t)}{\bar{y}}\right)^\omega\right)dt = \int_0^{\bar{y}} y\left(1-\left(\frac{y}{\bar{y}}\right)^\omega\right)dy = \frac{\omega\bar{y}^2}{2(2+\omega)}$$

and under the maintained assumption that almost all infections occur during the first year we conclude that $\psi^f = \psi 2(2+\omega)/(\omega\bar{y})$.

Suppose next that we replace ψ in (3) by $\psi(1-\ell y(t)/\bar{y})$, representing learning-by-doing. Consistent with evidence we assume that the learning effects let unit costs drop by a third over the course of the epidemic: $\ell = 1/3$.[69]

C.3 Numerical Methods

We use Mathematica to numerically solve for the value function and associated policy function. The domain is discretized using fourth-order finite difference methods to approximate the continuous HJB equation. The value function that solves the discretized HJB equation defined over the continuous state space converges to the value function of the original continuous HJB under our model assumptions (see Bardi and Capuzzo-Dolcetta (1997, theorem 1.1, section VI)). The method of lines is used for time integration and standard numerical methods for solving ODEs,

[68] The calibration of ψ is independent of our preference assumption. If we stipulate $r = 2$ rather than $r = 1$ (see footnote 66) then the annual utility costs of sacrificing the consumption share $1-\phi$ equals $365 \cdot \{-(a^\star)^{-1} + (a^\star\phi)^{-1}\} = -365(1-\phi^{-1})$. But $\ln(1-0.3581) \approx 1-(1-0.3071)^{-1}$, so the different value for r does not materially affect ψ.

[69] RECOVERY Collaborative Group et al. (2020) conducted a randomized trial in the UK finding that among 2104 patients that received dexamethasone mortality was reduced by one-third in patients receiving mechanical ventilation and by one-fifth in patients receiving oxygen but not mechanical ventilation.

such as the Runge-Kutta method, are used to solve for the discretized HJB. Barles and Souganidis (1991) show that the finite difference method satisfies the monotonicity, consistency and stability conditions that guarantee convergence of the approximation to the unique viscosity solution of the HJB equation.

As a robustness check we solve the HJB equation using the Initial Value Problem Differential-Algebraic Equations (IDA) method. This is a robust and efficient method designed for differential-algebraic equations that exploits the algebraic constraints in ODEs. The results we obtain are essentially identical.

D Viscosity Solutions of Non-Linear Partial Differential Equations

A generic HJB equation takes the form $v_d(y,d) + F(y, v(y,d), D_y v(y,d)) = 0$ where $v : \Omega \times [0,T] \to \mathbb{R}$ is a continuous value function; $\Omega \subseteq \mathbb{R}^n$ and $[0,T]$ denote the endogenous and exogenous state spaces, respectively; $D_y v(y,d)$ denotes the gradient with respect to the endogenous state; and F is smooth. Crandall and Lions (1983) introduce the notion of viscosity solution of a partial differential equation such as this HJB equation; for a textbook treatment see, e.g., Bardi and Capuzzo-Dolcetta (1997).

For all $(y,d) \in \Omega \times [0,T]$ define the superdifferential and subdifferential, respectively, of v as the following sets:

$$\begin{aligned} D^+ v(y,d) &= \left\{ p \in \mathbb{R}^n : \limsup_{x \to y} \frac{v(x,d) - v(y,d) - p.(x-y)}{|x-y|} \leq 0 \right\}, \\ D^- v(y,d) &= \left\{ q \in \mathbb{R}^n : \liminf_{x \to y} \frac{v(x,d) - v(y,d) - q.(x-y)}{|x-y|} \geq 0 \right\}. \end{aligned}$$

Note that if both $D^+ v(y,d)$ and $D^- v(y,d)$ are non-empty then $D^+ v(y,d) = D^- v(y,d)$ and v is differentiable at (y,d).

A continuous function w is a viscosity subsolution of the HJB equation if

$$w_d(y,d) + F(y, w(y,d), p) \leq 0 \; \forall p \in D^+ w(y,d), \;\; \forall (y,d) \in \Omega \times [0,T].$$

It is a viscosity supersolution if

$$w_d(y,d) + F(y, w(y,d), q) \geq 0 \; \forall q \in D^- w(y,d), \;\; \forall (y,d) \in \Omega \times [0,T].$$

If w is a viscosity subsolution and supersolution of the HJB equation then w is a viscosity solution of the HJB equation.

E Other Properties of the Value Function

Lemma 2. Under assumptions 1 and 2 and if $T = \infty$, V has a unique minimum at $y^{\min} < \bar{y}/(1+\omega)^{1/\omega}$. Parameter changes that imply a higher (lower) $y^{\min}$ also imply less (more) pronounced convexity of V at $y^{\min}$. Moreover,

$$\max_{y\in[y^{\min},\bar{y}]} V'(y) = V'(\bar{y}) = \frac{g(a^\star)\beta\bar{y}\omega\psi}{\rho+\nu+g(a^\star)\beta\bar{y}\omega} < \psi$$

and V is strictly convex over the domain $[y^{\min}, \bar{y}]$.

Proof. We assume that V is twice differentiable. From the government's HJB equation, the envelope condition with $T = \infty$ reads

$$\begin{aligned}(\rho+\nu)V'(y) &= -g(a(y))\beta\bar{y}\times \\ &\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y)) - y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right] \end{aligned} \tag{15}$$

Let $\hat{y}$ denote a point where V reaches a local minimum or maximum. Evaluated at $\hat{y}$ the envelope condition reduces to

$$\left(1-(1+\omega)\left(\frac{\hat{y}}{\bar{y}}\right)^{\omega}\right)\psi = \hat{y}\left(1-\left(\frac{\hat{y}}{\bar{y}}\right)^{\omega}\right)V''(\hat{y}). \tag{16}$$

Note that any extremum is locally unique since $V''(\hat{y}) = 0$ would only be consistent with $\hat{y} = \bar{y}/(1+\omega)^{1/\omega}$.

Uniqueness of $y^{\min}$. To see that V has a unique minimum suppose to the contrary that there exist multiple local minima. Consider two neighboring minima at, say, y^a and y^c with $y^a < y^c$. Then there must exist a local maximum at some y^b with $y^a < y^b < y^c$. Since $V''(y^a) > 0$, $V''(y^b) < 0$, and $V''(y^c) > 0$ the right-hand side of condition (16) evaluated at y^a, y^b, and y^c is strictly positive, negative, and positive, respectively. The sign of the left-hand side of condition (16) evaluated at the same points cannot alternate in this way. We have thus arrived at a contradiction which proves that V has a unique minimum, $y^{\min}$.

Upper Bound on $y^{\min}$. To derive a contradiction suppose first that $y^{\min} > \bar{y}/(1+\omega)^{1/\omega}$. Since the minimum is unique this implies $V'(\bar{y}/(1+\omega)^{1/\omega}) < 0$.

From the envelope condition (15),

$$(\rho+\nu)V'(\bar{y}/(1+\omega)^{1/\omega}) = g(a(\bar{y}/(1+\omega)^{1/\omega}))\beta\bar{y}^2 \frac{\omega}{(1+\omega)^{1+1/\omega}} V''(\bar{y}/(1+\omega)^{1/\omega})$$

and thus $V''(\bar{y}/(1+\omega)^{1/\omega}) < 0$. Since by assumption $y^{\min} > \bar{y}/(1+\omega)^{1/\omega}$ there must exist an inflection point, say y^i, with $\bar{y}/(1+\omega)^{1/\omega} < y^i < y^{\min}$, $V'(y^i) < 0$, and $V''(y^i) = 0$. But evaluated at y^i the envelope condition implies

$$\underbrace{(\rho+\nu)V'(y^i)}_{<0} = -\underbrace{g(a(y^i))\beta\bar{y}}_{>0} \underbrace{\left(1-(1+\omega)\left(\frac{y^i}{\bar{y}}\right)^{\omega}\right)}_{<0} \underbrace{(\psi - V'(y^i))}_{>0},$$

which yields a contradiction. We conclude that $y^{\min} \leq \bar{y}/(1+\omega)^{1/\omega}$.

In fact, $y^{\min} < \bar{y}/(1+\omega)^{1/\omega}$; for if $y^{\min}$ equalled $\bar{y}/(1+\omega)^{1/\omega}$ then the minimum could not be unique since condition (16) would imply $V''(\bar{y}/(1+\omega)^{1/\omega}) = 0$. We conclude that $y^{\min} < \bar{y}/(1+\omega)^{1/\omega}$.

Effect of Parameters on Convexity of V at $y^{\min}$. From equation (16), $V''(y^{\min}) = \left(1-(1+\omega)\left(\frac{y^{\min}}{\bar{y}}\right)^{\omega}\right)\psi/y^{\min}/\left(1-\left(\frac{y^{\min}}{\bar{y}}\right)^{\omega}\right)$. This is strictly decreasing in $y^{\min}$. The result then follows.

Maximum Slope of V. Let $\mathcal{Y} = [y^{\min}, \bar{y}]$ and suppose first that there exists no $y \in \mathcal{Y}$ such that $V''(y) = 0$ (no inflection point on the domain $\mathcal{Y}$). Since $V(y)$ has a unique global minimum at $y^{\min}$, $V(y)$ is strictly increasing and convex in this case for all $y \in \mathcal{Y} \setminus y^{\min}$, and thus $\max_{y\in\mathcal{Y}} V'(y) = V'(\bar{y})$. Suppose next that there exists some $y \in \mathcal{Y}$ such that $V''(y) = 0$ (at least one inflection point on the domain $\mathcal{Y}$). Let $\mathcal{Y}^i \subset \mathcal{Y}$ denote the set of inflection points and let $y^i = \arg\max_{y\in\mathcal{Y}^i} V'(y)$. Then, $\max_{y\in\mathcal{Y}} V'(y) = \max[V'(y^i), V'(\bar{y})]$. From the envelope condition (15) and the fact that $V''(y^i) = 0$, $(\rho+\nu)V'(y^i) = -g(a(y^i))\beta\bar{y}\left(1-(1+\omega)\left(\frac{y^i}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y^i))$ or

$$V'(y^i) = \frac{g(a(y^i))\beta\bar{y}\left((1+\omega)\left(\frac{y^i}{\bar{y}}\right)^{\omega}-1\right)\psi}{(\rho+\nu+g(a(y^i))\beta\bar{y}\left((1+\omega)\left(\frac{y^i}{\bar{y}}\right)^{\omega}-1\right)} \leq \frac{g(a^\star)\beta\bar{y}\omega\psi}{\rho+\nu+g(a^\star)\beta\bar{y}\omega} = V'(\bar{y}),$$

where the weak inequality follows from $(1+\omega)\left(\frac{y^i}{\bar{y}}\right)^{\omega} - 1 \leq \omega$, $a(y^i) \leq a^\star$, and the fact that g is increasing; and the equality on the right-hand side follows directly from condition (15). Accordingly, $\max_{y\in\mathcal{Y}} V'(y) = V'(\bar{y})$. We conclude that in either case $\max_{y\in\mathcal{Y}} V'(y) = V'(\bar{y})$.

Strict Convexity of V on the Domain $\mathcal{Y}$. From the previous results $V(y)$ is weakly convex in a neighborhood of $\bar{y}$ and strictly convex at $y^{\min}$.

[1] Suppose first that there do not exist open intervals on the domain $\mathcal{Y}$ such that $V''(y) = 0$ for all points in the interval (however, there may exist points $y \in \mathcal{Y}$ with $V''(y) = 0$). Given the convexity of V at $\bar{y}$ and $y^{\min}$, the number of inflection points on the domain $\mathcal{Y}$ at which $V'(y)$ changes signs must be even (including equal to zero). To arrive at a contradiction suppose that the number is strictly positive and consider the smallest two neighboring inflection points on the domain $\mathcal{Y}$, say, y^a and y^b with $y^a < y^b$. Given the specified assumptions, $0 < V'(y^a) < \psi$ (using the above argument on the maximum slope of V), $V''(y^a) = V''(y^b) = 0$, and $V''(y) < 0 < V'(y)$ for all $y \in (y^a, y^b)$. Note that $y^a > \bar{y}/(1+\omega)^{1/\omega}$ since the envelope condition (15) and $V''(y^a) = 0$ imply $(\rho+\nu)V'(y^a) = -g(a(y^a))\beta\bar{y}\left(1-(1+\omega)\left(\frac{y^a}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y^a))$ such that $0 < V'(y^a) < \psi$ requires $1-(1+\omega)\left(\frac{y^a}{\bar{y}}\right)^{\omega} < 0$. Accordingly, $1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega} < 0$ for all $y \in (y^a, y^b)$.

Rewriting the envelope condition (15) as

$$g(a(y)) = -\frac{(\rho+\nu)V'(y)}{\beta\bar{y}\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y)) - y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right]} \tag{17}$$

and differentiating with respect to y gives, for all $y \in (y^a, y^b)$,

$$\begin{aligned} g'(a(y))a'(y) &= -\frac{(\rho+\nu)V''(y)}{\beta\bar{y}\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y)) - y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right]} \\ &\quad + \frac{(\rho+\nu)V'(y)D}{\beta\bar{y}\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y)) - y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right]^2} \\ &= \underbrace{\frac{V''(y)}{V'(y)}g(a(y))}_{<0} + \underbrace{(\rho+\nu)V'(y)}_{>0}\frac{-\{>0\} - 2\{>0\} - \{>0\}V'''(y)}{\{>0\}}, \end{aligned}$$

where $D \equiv (V'(y)-\psi)\omega(1+\omega)\frac{y^{\omega-1}}{\bar{y}^{\omega}} - 2\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y) - y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V'''(y)$ and $\{>0\}$ denotes strictly positive terms (which might differ from each other). Since V is strictly concave for $y \in (y^a, y^b)$ there exists some nonempty $\mathcal{Z} \subseteq (y^a, y^b)$ with $V'''(y) > 0$ for all $y \in \mathcal{Z}$. Since g is increasing the preceding equality implies that $a'(y) < 0$ for all $y \in \mathcal{Z}$.

The government's first-order condition implies

$$\frac{u'(a(y))}{g'(a(y))} = \beta\bar{y}y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi - V'(y)). \tag{18}$$

Differentiating the right-hand side with respect to y yields

$$\beta\bar{y}\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi-V'(y))-y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right],$$

which is strictly negative for all $y \in (y^a, y^b)$ because of condition (17) and the fact that $g(a(y)) > 0$. Accordingly, the right-hand side of equation (18) is strictly decreasing in y, and so must be the left-hand side. This requires $a'(y) > 0$ for all $y \in (y^a, y^b)$ since g is weakly convex (assumption 1) and u strictly concave (assumption 2). In particular, $a'(y) > 0$ for all $y \in \mathcal{Z}$. We have thus arrived at a contradiction and conclude that there exist no inflection points on the domain $\mathcal{Y}$.

[2] Suppose next that there do exist open intervals on the domain $\mathcal{Y}$ such that $V''(y) = 0$ for all points in the interval (there may also exist inflection points). Consider such an open interval, say, (y^a, y^b). Then, $V''(y) = 0$ for all $y \in (y^a, y^b)$. Differentiating the envelope condition with respect to y gives, for all $y \in (y^a, y^b)$,

$$\begin{aligned}
g'(a(y))a'(y) &= -\frac{(\rho+\nu)V''(y)}{\beta\bar{y}\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi-V'(y))-y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right]} \\
&\quad+\frac{(\rho+\nu)V'(y)D}{\beta\bar{y}\left[\left(1-(1+\omega)\left(\frac{y}{\bar{y}}\right)^{\omega}\right)(\psi-V'(y))-y\left(1-\left(\frac{y}{\bar{y}}\right)^{\omega}\right)V''(y)\right]^2} \\
&= \underbrace{(\rho+\nu)V'(y)}_{>0}\frac{(V'(y)-\psi)\omega(1+\omega)\frac{y^{\omega-1}}{\bar{y}^{\omega}}}{\{>0\}},
\end{aligned}$$

where $\{> 0\}$ denotes a strictly positive term and where we use the fact that $V''(y) = V'''(y) = 0$ for all $y \in (y^a, y^b)$. Since g is increasing the preceding equality implies that $a'(y) < 0$ for all $y \in (y^a, y^b)$.

Differentiating the right-hand side of the government's first-order condition with respect to y yields a strictly negative expression for all $y \in (y^a, y^b)$, because of condition (17), $g(a(y)) > 0$, and $V''(y) = 0$. Accordingly, the right-hand side of equation (18) is strictly decreasing in y, and so must be the left-hand side, which requires $a'(y) > 0$ for all $y \in (y^a, y^b)$. We have thus arrived at a contradiction and conclude that there exist no open intervals on the domain $\mathcal{Y}$ such that $V''(y) = 0$ for all points in the interval.

[3] We conclude from [1] and [2] that V is strictly convex on the domain $\mathcal{Y}$. □