SUPPLEMENT TO
"OPTIMAL TAXATION AND R&D POLICIES"

by Ufuk Akcigit, Douglas Hanley, and Stefanie Stantcheva

S.1 Optimal Policies in a Simple Two-type, One-Period Model

In this section, we illustrate the underlying logic of the optimal mechanism in a very simple two-type, one-period model.

Suppose that firms can be of the high research productivity type θ_2 or of the low productivity type θ_1. The fractions in the population of firms of types high and low are, respectively, f_2 and f_1, with $f_2 = 1 - f_1$. The problem is static: Firms enter period 1 with a knowledge of their type realization, chose R&D investments $r(\theta_i)$ and R&D effort $l(\theta_i)$ at the beginning of the period. The step size is $\lambda(\theta_i) = \lambda(r(\theta_i), l(\theta_i), \theta_i)$ and quality is $q(\theta_i) = q_0 + \lambda(\theta_i)$, where q_0 is given. At the end of the period firms receive a transfer $T(\theta_i)$ from the government. For the exposition, suppose that the step size takes the form:

$$\lambda(r, l, \theta_i) = w(r, \theta_i) l$$

for an increasing and concave function w. The market structure between the intermediate goods and the final goods producer generates a demand function $p(q, k)$ for the intermediate goods. With full patent protection in place, the intermediate good producer faces the monopolist price. Profits are denoted by $\pi(q, \bar{q})$ as a function of quality q and aggregate quality $\bar{q} = f_1 q(\theta_1) + f_2 q(\theta_2)$.

In the planning problem, the planner sets a menu of contracts $(r(\theta_i), l(\theta_i), T(\theta_i))$ for $i = 1, 2$ and lets firms self-select allocations from this menu. For simplicity, we set $\chi = 1$.[1] For any quality, the firm will choose the privately optimal quantity, leading to output net of production costs $\tilde{Y}(q(\theta_i), \bar{q})$ for type θ_i. The remaining components of the menu $(r(\theta_i), l(\theta_i), T(\theta_i))_{i=1,2}$ and $\bar{q}$ are chosen to maximize social welfare defined in (3), and which in this simple case becomes:

$$W = f_1 \left(\tilde{Y}(q(\theta_1), \bar{q}) - M(r(\theta_1)) - T(\theta_1) \right) + f_2 \left(\tilde{Y}(q(\theta_2), \bar{q}) - M(r(\theta_2)) - T(\theta_2) \right),$$

subject to $q(\theta_i) = q_0 + \lambda(\theta_i)$ with q_0 given, and subject to firms' participation constraints:

$$T(\theta_i) - \phi(l(\theta_i)) \geq 0.$$

We can also allow for some different thresholds in the participation constraint, such that $T(\theta_i) - \phi(l(\theta_i)) \geq \underline{V}(\theta_i)$. In the first best, firm type is observable, $\chi = 0$, and the planner makes each firm invest the efficient level of effort and inputs, such that the marginal effort and R&D investment costs equal the social impact, as in section 2.3, and surplus is extracted in a lump-sum fashion from the firms, i.e.,[2]

$$T(\theta_i) = \phi(l(\theta_i)).$$

[1] This is without loss of generality: a $\chi \neq 1$ would simply appear as a scaling factor in front of the screening term in the formulas below.

[2] More precisely,

$$M'(r(\theta_i)) = \left(\frac{\partial \tilde{Y}(q(\theta_i), \bar{q})}{\partial q} + \left(f_1 \frac{\partial \tilde{Y}(q(\theta_1), \bar{q})}{\partial \bar{q}} + f_2 \frac{\partial \tilde{Y}(q(\theta_2), \bar{q})}{\partial \bar{q}} \right) \right) \frac{\partial \lambda(r(\theta_i), l(\theta_i), \theta_i)}{\partial r(\theta_i)}$$

$$\frac{\phi(l(\theta_i))}{w(r(\theta_i), \theta_i)} = \frac{\partial \tilde{Y}(q(\theta_i), \bar{q})}{\partial q} + \left(f_1 \frac{\partial \tilde{Y}(q(\theta_1), \bar{q})}{\partial \bar{q}} + f_2 \frac{\partial \tilde{Y}^*(q(\theta_2), \bar{q})}{\partial \bar{q}} \right)$$

The second-best problem imposes an incentive constraint for each type i:

$$T(\theta_i) - \phi(l(\theta_i)) \geq T(\theta_j) - \phi\left(\frac{w(r(\theta_j),\theta_j)l(\theta_j)}{w(r(\theta_j),\theta_i)}\right) \quad \forall (i,j).$$

Given that the goal is to minimize total transfers to the firms, one can show that the incentive constraint of type θ_2 and the participation constraint of type θ_1 will be binding.[3] Indeed, at the first-best allocations and transfer levels, high research productivity firms will be tempted to pretend that they are low productivity firms. This is because they have to forfeit all their surplus to the planner, but, since they are able to reach any step size at a lower R&D effort cost than low research productivity firms, they could achieve a positive surplus by selecting the low research productivity firm's first-best allocation. To prevent this from happening, the allocation of the low research productivity firms needs to be distorted so as to make it less attractive to high productivity firms.

The transfers then have to satisfy:

$$T(\theta_1) = \phi\left(l(\theta_1)\right)$$

$$T(\theta_2) - \phi\left(l(\theta_2)\right) \geq T(\theta_1) - \phi\left(\frac{w(r(\theta_1),\theta_1)l(\theta_1)}{w(r(\theta_1),\theta_2)}\right).$$

Substituting these expressions into the social objective, we obtain the so-called virtual surplus, which is social surplus minus the informational rent forfeited to the high type θ_2 to induce him to truthfully reveal his type. The social optimum will maximize allocative efficiency (the first line below) while trying to reduce the informational rent forfeited to the high type (the second line):

$$\begin{aligned} W = f_1\left(\tilde{Y}(q_1(\theta_1),\bar{q}) - M(r(\theta_1)) - \phi\left(l(\theta_1)\right)\right) + f_2\left(\tilde{Y}(q(\theta_2),\bar{q}) - M(r(\theta_2) - \phi\left(l(\theta_2)\right))\right) \\ - f_2\left(\phi\left(l(\theta_1)\right) - \phi\left(\frac{w(r(\theta_1),\theta_1)l(\theta_1)}{w(r(\theta_1),\theta_2)}\right)\right). \qquad \text{(S1)} \end{aligned}$$

Characterization of the Optimal Allocation in Terms of Wedges. The constrained efficient allocation can be described using so-called wedges or implicit taxes and subsidies, which measure the deviation of the allocation relative to the laissez-faire economy with patent protection. In the laissez-faire economy with patent protection, profits are a function of the product's quality and aggregate quality, $\pi(q(\theta_i),\bar{q})$, as defined in Section 2. The effort wedge, $\tau(\theta_i)$ on type θ_i is defined as the gap between the marginal *private* benefit of effort and its cost, while the R&D investment wedge is defined as the gap between the marginal cost of R&D and its marginal private benefit. Thus, a higher effort wedge means a lower incentive for R&D effort, while a higher R&D investment wedge means a higher incentive for R&D investments. Formally:

$$s(\theta_i) = M'(r(\theta_i)) - \frac{\partial \pi(q(\theta_i),\bar{q})}{\partial q(\theta_i)} \frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial r(\theta_i)}$$

$$(1-\tau(\theta_i))\frac{\partial \pi(q(\theta_i),\bar{q})}{\partial q(\theta_i)} \frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial l(\theta_i)} = \phi'(l(\theta_i)).$$

In the implementation below, it will be clear that there is a very natural map between the wedges (i.e., implicit taxes and subsidies) and the explicit marginal tax rates of the implementing tax function.

[3] As is usual in these types of screening problems, the slackness of the low type's omitted incentive constraint can be checked ex post.

Taking the first-order conditions of the social objective with respect to $r(\theta_i)$ and $l(\theta_i)$ for $i = 1, 2$ and using the definitions of the wedges, we obtain that for the low research productivity type, the allocations are distorted just enough to balance the informational rent forfeited to the high type and the loss in allocative efficiency.

Proposition 1. ***Optimal Allocations for Low Research Productivity Firms.***
i) The optimal R&D investment wedge on the low research productivity type is given by:

$$s(\theta_1) = \underbrace{\left(f_1 \frac{\partial \tilde{Y}(q(\theta_1), \bar{q})}{\partial \bar{q}} + f_2 \frac{\partial \tilde{Y}(q(\theta_2), \bar{q})}{\partial \bar{q}}\right) \frac{\partial w(r(\theta_1), \theta_1)}{\partial r} l(\theta_1)}_{\text{Pigouvian corection}}$$
$$+ \underbrace{\left(\frac{\partial \tilde{Y}(q(\theta_1), \bar{q})}{\partial q(\theta_1)} - \frac{\partial \pi(q(\theta_1), \bar{q})}{\partial q(\theta_1)}\right) \frac{\partial w(r(\theta_1), \theta_1)}{\partial r} l(\theta_1)}_{\text{Monopoly quality valuation correction}}$$
$$\underbrace{+ \frac{f_2}{f_1} \underbrace{\left(1 - \frac{\frac{\partial \log(w(r(\theta_1), \theta_2))}{\partial \log(r)}}{\frac{\partial \log(w(r(\theta_1), \theta_1))}{\partial \log(r)}}\right)}_{\text{Complementarity}} \frac{\frac{\partial w(r(\theta_1), \theta_1)}{\partial r} l(\theta_1)}{w(r(\theta_1), \theta_2)} \phi' \left(\frac{w(r(\theta_1), \theta_1) l(\theta_1)}{w(r(\theta_1), \theta_2)}\right)}_{\text{Screening term}}. \tag{S2}$$

ii) The optimal R&D effort wedge on the low productivity firm is given by:

$$\tau(\theta_1) \frac{\partial \pi(q(\theta_1), \bar{q})}{\partial q(\theta_1)} = -\left(\frac{\partial \tilde{Y}(q(\theta_1), \bar{q})}{\partial q(\theta_1)} - \frac{\partial \pi(q(\theta_1), \bar{q})}{\partial q(\theta_1)}\right) - \left(f_1 \frac{\partial \tilde{Y}(q(\theta_1), \bar{q})}{\partial \bar{q}} + f_2 \frac{\partial \tilde{Y}(q(\theta_2), \bar{q})}{\partial \bar{q}}\right)$$
$$+ \underbrace{\frac{f_2}{f_1} \left(\frac{1}{w(r(\theta_1), \theta_1)} \phi'(l(\theta_1)) - \frac{1}{w(r(\theta_1), \theta_2)} \phi' \left(\frac{w(r(\theta_1), \theta_1) l(\theta_1)}{w(r(\theta_1), \theta_2)}\right)\right)}_{\text{Screening term: Cost differential between high and low productivity firms}}. \tag{S3}$$

Proof. Taking the first-order conditions of the planner's problem in (S1) with respect to $l(\theta_i)$ and $r(\theta_i)$ for each $i = 1, 2$ and using the definitions of the wedges yields the formulas. □

The optimal implicit subsidy on R&D investment in (S2) and the R&D effort wedge in (S3) balance three considerations.

1) Pigouvian correction for technology spillovers: Incentives are increasing in the Pigouvian correction that aligns private incentives with the social benefit from R&D technology spillovers, which are the key reason for the government to intervene. This correction is larger when the marginal return to R&D investments $\left(\frac{\partial w(r(\theta_1), \theta_1)}{\partial r}\right)$ is larger.

2) Monopoly quality valuation correction: Starting from a laissez-faire with patent protection, the monopolist values each marginal increase in quality less than its marginal social value: this difference in quality valuation must also be corrected for in the optimal planning problem. This is the second term in each of the wedge formulas. The distortions in the R&D investment and effort are modified so as to indirectly compensate for the under-provision of quantity of the monopolist. The effect of a change in quantity (induced by extra investment in R&D investment or R&D effort) on social welfare, implicit in $\frac{\partial \tilde{Y}(q(\theta_i), \bar{q})}{\partial q(\theta)}$, is first-order and is proportional to the monopoly distortion, i.e., the gap between price and marginal cost.[4] The pre-existing monopoly

[4]Formally, $\frac{\partial \tilde{Y}(q(\theta_i), \bar{q})}{\partial q(\theta_i)} = \frac{\partial Y(q(\theta_i), k(q(\theta_i), \bar{q}))}{\partial q(\theta_i)} + \left(p(q(\theta_i), k(q(\theta_i), \bar{q})) - \frac{\partial C(k(q(\theta_i), \bar{q}), \bar{q})}{\partial k}\right) \frac{\partial k(q(\theta_i), \bar{q})}{\partial q(\theta_i)}$ where $k(q(\theta_i), \bar{q})$ is the quantity chosen to maximize profits by a monopolist with quality $q(\theta_i)$.

distortions amplify the direct impact of R&D effort and investment on output and the indirect impact through the technology spillover, pushing the R&D effort wedge down and the R&D investment wedge up. The optimal R&D policies hence depend on the IPR policies in place. If there was no monopoly distortion in the laissez-faire economy, i.e., if there was for instance a prize system, then there would be no need to correct for it and this term would disappear from the optimal wedge formulas.[5]

3) Screening term: The screening term (the third term in each formula) captures the modification to the first-best incentive that is induced by the asymmetric information. It is decreasing in the fraction of high research productivity firms over low research productivity firms: the lower the fraction of low productivity firms, and the less costly it is to distort their effort or investments for the sake of reducing the informational rent of the (more frequently encountered) high productivity firms.

The screening term depends on the relative complementarity of R&D investments with R&D effort versus firm research productivity. Since the step size is assumed here to be multiplicatively separable, the elasticity of the step size to R&D effort for both types is just 1, the first term in the "complementarity" term. The relative elasticity of the return to effort $w(r,\theta)$ with respect to R&D for the high and the low type, $\frac{\partial \log(w(r(\theta_1),\theta_2))}{\partial \log(r)} / \frac{\partial \log(w(r(\theta_1),\theta_1))}{\partial \log(r)}$ measures how complementary R&D investments are to firm research productivity: if the elasticity is increasing in type, then R&D investments benefit disproportionately high research productivity firms. The more elastic the high type's return is to R&D, the less the R&D investment of the low type can be subsidized, as this makes it more tempting for the high type to pretend to be low type. Put differently, increasing R&D investments of the low type when the relative elasticity is high means tightening the high type's incentive constraint and giving that firm more informational rent. As a special case, if the elasticities of the high and low types are the same, then R&D investments of the low type do not affect the high type's incentive constraint. As a result, the screening term drops out and the optimal marginal R&D subsidy is set solely to correct for the technology spillover and the monopoly distortion.

Stimulating R&D investments is beneficial when there is a high complementarity of R&D investments with unobservable R&D effort, because it stimulates the unobservable input, but is detrimental when there is a high complementarity with firm research productivity, as it then tightens the incentive constraint of the high research productivity firm. The basic logic is that investments in R&D are distorted only in so far as they (beneficially) affect the incentive constraint of the high research productivity firm, i.e., as long as they can indirectly stimulate the unobservable R&D effort choice.

For the R&D effort wedge, the efficiency cost of distorting the low research productivity firm's R&D efforts depends on the comparative productive advantage of the high type relative to the low type. The efficiency cost depends on the difference in the marginal cost $\phi'(l)$ of producing the step size assigned to the low research productivity firm (which is $\lambda(\theta_1)$) between the low and the high research productivity firm. Since the cost function $\phi(l)$ is convex, this difference is always positive. The smaller this difference, the more tempting it is for the high research productivity firm to imitate the low research productivity one and the more the R&D effort of low productivity firms should be reduced. This increases the optimal effort wedge $\tau(\theta_1)$ on the low productivity firm's R&D effort.

On the other hand, the high research productivity firms' allocations are set based on the

[5]Naturally, larger wedges (i.e., distortions relative to the laissez-faire) do not imply in any sense that there is more investment in effort or R&D relative to a situation with smaller wedges.

monopoly valuation and Pigouvian correction terms only. The screening term is zero since the low type's incentive constraint is not binding. Section S.1 explains two possible implementations of the optimal allocations in this simple model and provides expressions for the marginal tax rates and the marginal subsidy rate in the case in which this implementing tax system can be made differentiable.

R&D Policies when Production can be Controlled. Imagine now that the government can also intervene in the private market between intermediate and final goods producers and make the policies contingent on the quantity produced. As a result, for any quality, the socially optimal quantity can be enforced and output net of production costs is $\tilde{Y}^*(q(\theta_i),\bar{q})$ for type θ_i. This is because the optimal quantity to be produced is only conditional on quality and there is no reason to distort it (although the quality decision itself will still be distorted relative to the first best). The planning problem, and hence the optimal wedges, are the same, but with $\tilde{Y}^*(q(\theta_i),\bar{q})$ replacing $\tilde{Y}(q(\theta_i),\bar{q})$ in (S2) and (S3). Since the optimal quantity can now be implemented, the value of each incremental quality improvement is even larger (relative to private firm profits) and it is optimal to foster innovation even more with larger R&D wedges and lower corporate wedges. Another way of putting this is that when quantity can be controlled, the planner will optimally make the firm deviate even more from the allocation it would have picked in the laissez-faire.

S.1 Implementation

We illustrate here the two implementations in the case in which quantity can also be controlled. The benchmark case where quantity cannot be controlled is treated in detail in Section 4.2.

Tax Implementation. First, the government can subsidize the price of production at a nonlinear rate $s_p(k,q)$ as a function of the quantity and quality of the good sold to the final good producer, such that the post subsidy price is $(1+s_p(k,q))p(k,q) = \frac{Y(k,q)}{k}$, and in addition levy a profit tax (which could be negative) $T(\pi,r)$ that depends nonlinearly on profits and R&D investments. Firms choose quantity to maximize profits conditional on quality, which, thanks to the price subsidy, becomes equivalent to maximizing household consumption net of production costs. Note that under a constant monopoly price markup (as arises for instance under the functional form assumptions in Section 5 where $Y(q,k) = \frac{1}{1-\beta}q^{\beta}k^{1-\beta}$), the price subsidy needed to align the monopolist's post-tax price with social marginal valuation of quantity is constant and equal to $\frac{\beta}{1-\beta}$. With this price subsidy, profits will be equal to $\tilde{Y}^*(q_0+\lambda(r,l,\theta_i),\bar{q})$. The maximization problem of a firm of type θ_i with respect to the remaining choices of l and r is then:

$$\max_{l,r}\{\tilde{Y}^*(q_0+\lambda(r,l,\theta_i),\bar{q}) - T(\tilde{Y}^*(q_0+\lambda(r,l,\theta_i),\bar{q}),r) - \phi(l) - M(r)\}.$$

The first-order conditions of the firm with this tax implementation are:

$$\begin{aligned}&-\frac{\partial T(\tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial r(\theta_i)}\\&+\frac{\partial \tilde{Y}^*(q(\theta_i),\bar{q})}{\partial q}\frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial r(\theta_i)}\left(1-\frac{\partial T(\tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial \pi}\right) = M'(r(\theta_i))\\&\left(1-\frac{\partial T(\tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial \pi}\right)\frac{\partial \tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial q}\frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial l(\theta_i)} = \phi'(l(\theta_i)).\end{aligned}$$

We can use the first-order conditions of the firms into the optimal wedge formulas to obtain a characterization of the optimal (explicit) marginal tax and subsidy:

$$-\frac{1}{\frac{\partial w(r(\theta_1),\theta_1)}{\partial r}l(\theta_1)}\frac{\partial T(\tilde{Y}^*(q(\theta_1),\bar{q}),r(\theta_1))}{\partial r(\theta_1)} = \frac{\partial T(\tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial \pi}\frac{\partial \tilde{Y}^*(q(\theta_i),\bar{q})}{\partial q}$$
$$+\left(f_1\frac{\partial \tilde{Y}^*(q(\theta_1),\bar{q})}{\partial \bar{q}}+f_2\frac{\partial \tilde{Y}^*(q(\theta_2),\bar{q})}{\partial \bar{q}}\right)+\frac{f_2}{f_1}\left(1-\frac{\frac{\partial \log(w(r(\theta_1),\theta_2))}{\partial \log(r)}}{\frac{\partial \log(w(r(\theta_1),\theta_1))}{\partial \log(r)}}\right)\frac{1}{w(r(\theta_1),\theta_2)}\phi'\left(\frac{w(r(\theta_1),\theta_1)l(\theta_1)}{w(r(\theta_1),\theta_2)}\right)$$

$$\frac{\partial T(\tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial \pi}\frac{\partial \tilde{Y}^*(q(\theta_i),\bar{q}),r(\theta_i))}{\partial q} = -\left(f_1\frac{\partial \tilde{Y}^*(q(\theta_1),\bar{q})}{\partial \bar{q}}+f_2\frac{\partial \tilde{Y}^*(q(\theta_2),\bar{q})}{\partial \bar{q}}\right)$$
$$-\frac{f_2}{f_1}\left(\frac{1}{w(r(\theta_1),\theta_2)}\phi'\left(\frac{w(r(\theta_1),\theta_1)l(\theta_1)}{w(r(\theta_1),\theta_2)}\right)-\frac{1}{w(r(\theta_1),\theta_1)}\phi'(l(\theta_1))\right).$$

Note that the monopoly quality valuation correction term does not enter the optimal tax and subsidy because the monopoly quantity distortion is taken care of by the price subsidy in this implementation. The profits that the firm maximizes are exactly equivalent to $\tilde{Y}^*$, the socially valued output net of production costs.

Implementation with a Prize Mechanism. The government can also simply purchase the innovation directly from the firm in exchange for a prize $G(\lambda, r)$ that depends on the step size (or, interchangeably, on the realized quality q) and on R&D investment. If the prize function is differentiable in its two arguments, the formulas for the marginal change in prize with respect to the step size or R&D investments can immediately be obtained by substituting for the wedges in the planner's first-order conditions, using the link between the wedges and the marginal prize with respect to product quality and R&D expenses.

$$s(\theta_i) = \frac{\partial G(\lambda(r(\theta_i),l(\theta_i),\theta_i),r(\theta_i))}{\partial r(\theta_i)}+\frac{\partial G(\lambda(r(\theta_i),l(\theta_i),\theta_i),r(\theta_i))}{\partial \lambda}\frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial r(\theta_i)}$$
$$\tau(\theta_i)\frac{\partial \pi(q(\theta_i),\bar{q})}{\partial q(\theta_i)}\frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial l(\theta_i)} = \frac{\partial G(\lambda(r(\theta_i),l(\theta_i),\theta_i),r(\theta_i))}{\partial \lambda}\frac{\partial \lambda(r(\theta_i),l(\theta_i),\theta_i)}{\partial l(\theta_i)}.$$

S.2 Controlling Quantity

In the main part of the paper, the government is assumed to be unable to control the quantity of production. This means that the government takes the demand function and the price for each unit of quality as given. There is a tight link between this and the ability to control IPR policy. Not being able to control quantity produced essentially amounts to taking the patent system as given. The main paper thus considers what the optimal corporate tax and R&D policies should be in the presence of a patent system. Here, we consider the case in which the government is free to set the quantity of production conditional on quality, which means the government can in fact freely set the intellectual property policy as well, which in this case is a prize system through which the government directly purchases the innovation from the producer.

If the government were also able to intervene in the output market and control the quantity produced, the planning problem is identical to P in (8), except that the impact of quality improvements on the net output produced by the monopolist, $\tilde{Y}(q_t(\theta^s),\bar{q}_s)$, is replaced everywhere with the impact of quality improvement on net output as would optimally be chosen by the planner, for every quality level, i.e. $\tilde{Y}^*(q_t(\theta^s), \bar{q}_s)$. Accordingly, in the optimal wedge formulas in Proposition (1), $Q_{t+1}(\theta^{t+1})$ is replaced by $Q^*_{t+1}(\theta^{t+1})$. All else equal, when the planner is also able to control quantity, wedges are larger because the planner is able to make the firm deviate more relative to what it would (suboptimally) do in the laissez-faire. Because being able to control quantity implies removing a constraint in the planner's problem, total output net of all costs will be higher than when quantity cannot be controlled.

Implementation. The constrained efficient allocation from program $P(\bar{q})$ (when quantity is observable) can then be implemented in two ways, which from a theoretical point of view are equivalent. The first implementation features a price subsidy $s_p(k,q)$ such that the post-subsidy price perceived by the intermediate good producer is $p(k,q)(1+s_p(k,q)) = \frac{Y(k,q)}{k}$. In this case, the private producer will maximize profits equal to $Y(k,q) - C(k,\bar{q})$ conditional on q, which is exactly the social surplus from production k. This price subsidy should be combined with a comprehensive, age-dependent tax function $T_t(q_t, r_t, q_{t-1}, r_{t-1}, q_1)$ that conditions on current quality q_t, lagged quality, q_{t-1}, current R&D, r_t, lagged R&D r_{t-1}, and first-period quality q_1.

Second, the government could set up a prize mechanism, through which it purchases the new innovation flow (i.e., the step size) λ_t from the firm in each period, and produces the socially optimal quantity of the good of quality $q_t = (1-\delta)q_{t-1} + \lambda_t$. Here, the government becomes the central owner of the intellectual property and keeps adding to its stock every period, in exchange for a prize. The prize amount $G_t(\lambda_t, r_t, r_{t-1}, q_1)$ paid for an innovation λ_t depends on firm age, current and lagged R&D investments, and the initial quality q_1.

S.1 Numerical Simulation: Optimal Allocations and Wedges When Quantity can be Controlled

When quantity can also be controlled, the planner has an additional lever that can also be made part of the contract. As a result, the planner can make firms deviate even more from their laissez-faire allocations to induce a better allocation. Accordingly, the wedges are larger in absolute value, as illustrated in Figure S1. Overall, the innovation inputs and step sizes are larger, as shown in Figure S2.

S.2 Welfare Gains from Simpler Policies When Quantity can be Controlled

Table S.I shows the welfare gains from simpler policies relative to the optimal contract when quantity can be controlled. In this table, each panel considers a separate class of policies, ranging step-by-step from linear to nonlinear and non-separable ones. We show the welfare achieved from the optimal policy in each class relative to the planning problem in which quantity can be controlled. The first row shows the welfare level achieved by the current policies in the U.S., which are approximated with a linear 23% effective corporate tax rate and a 19% effective R&D subsidy rate.

S.3 Compustat Data Matched to Patent Data

In this section, we redo our analysis on the sample made of only publicly traded firms, based on COMPUSTAT data matched to patent data. For this purpose, we select our sample so as to make it as close as possible to the one in Bloom, Schankerman, and Van Reenen (2013). The

TABLE S.I: WELFARE FROM OPTIMAL SIMPLER POLICIES WHEN QUANTITY CAN BE CONTROLLED

Policy Type	**Welfare Achieved Relative to Full Optimum**
A. Current US policy	
$T'(\pi) = 0.23 \qquad S'(M) = 0.19$	7%
B. Optimal Linear	
$T'(\pi) = \tau_0 \qquad S'(M) = s_0$	92.4%
C. Linear with Interaction Term	
$T'(\pi, M) = \tau_0 + \tau_1 M \qquad S'(M) = s_0$	95.1%
D. Heathcote-Storesletten-Violante (HSV)	
$T'(\pi) = \tau_0 - \tau_1 \pi^{\tau_2} \qquad S'(M) = s_0 - s_1 M^{s_2}$	96.3%
E. HSV Tax on Profits and Linear Subsidy	
$T'(\pi) = \tau_0 \qquad S'(M) = s_0 - s_1 M^{s_2}$	95.8%
F. HSV Subsidy on R&D and Linear Profit Tax	
$T'(\pi) = \tau_0 \qquad S'(M) = s_0 - s_1 M^{s_2}$	96.2%
G. HSV with Interaction Term	
$T'(\pi, M) = \tau_0 + \tau_3 M^{s_2} - \tau_1 \pi^{\tau_2}$ $S'(M) = s_0 - s_1 M^{s_2}$	96.4 %

Notes: The table shows the share of welfare from the full unrestricted optimum when quantity can be controlled that is achieved by the optimal policy within each class. Each panel shows a different class.

sample selection procedure that follows Bloom, Schankerman, and Van Reenen (2013) keeps all firms who patent at least once since 1963, so that they can at least at some point be matched to the patent data (this is natural also in light of our theory, which focuses on innovating firms). The final unbalanced panel contains 736 firms that are observed at least four times in the period 1980 to 2001 and is essentially identical to the sample in Bloom, Schankerman, and Van Reenen (2013).[6] Table S.II provides some summary statistics from the data.

Table S.II: Summary Statistics in the Compustat and Patent Data

Variable	Mean	Median
Sales (in mil. USD)	3133	494
Citations per patent	7.7	6
Patents per year	18.5	1
R&D spending / sales	0.043	0.014
Number of employees (000's)	18.4	3.8
Number of firms	736	

Note: The sample is selected to match as closely as possible the one in Bloom, Schankerman, and Van Reenen (2013), who keep firms that patent at least once since 1963 and which are observed for at least four years between 1980 and 2001.

S.4 Policies with a Finite Firm Life Cycle

One reason for time-dependent policies that is not covered in the paper is if firms have a finite lifecycle, i.e., if the maximum age is $T < \infty$. This leads to life cycle considerations such as the shorter horizon for any investments made later in firms' lives. Here the relevant issue is the distance of the period under consideration to the final period T. Both the laissez-faire and the socially optimal investments would naturally decline over a firm's life-cycle, all else equal, as earlier investments contribute to research productivity for more periods. If the technology spillover is positive, as seems natural, the Pigouvian correction term is always positive and, all else constant, will decline over time as the horizon shortens. This age-driven channel is fully eliminated by letting the horizon go to infinity, as we do in our benchmark case. Here, we provide the optimal policies with a finite life cycle.

Figures S5 and S6 show what happens when the life cycle is finite, with a given death and exit rate. In this case, the age paths of optimal inputs are hump-shaped, driven by the balance of the screening considerations and the life cycle considerations. In the first part of the life cycle, the screening considerations dominate; in the latter part, the dominant forces are the finite life cycle and the approach of the terminal period, which make investments less lucrative, privately and socially. Thus, with a finite life cycle, young firms, up to mid-life, should optimally provide an increasing amount of effort and investments for R&D. After mid-life, the effort and investment are declining given the shortening horizon left to reap the benefits.

[6] The results are robust to this sample selection. We repeated the analysis on a much broader sample of 6,400 firms over the period 1976 to 2006 that could be matched to the patent data for any year (without restricting to firms that are observed for at least four years). The results on this alternative sample are similar and are available upon demand.

S.5 Robustness Checks on Parameters and Moments

We provide here robustness checks and sensitivity analyses for our estimation.

In Figures S7 and S8, we perform a type of two-step GMM estimation with weights taken from the variance-covariance matrix of moments. The reason this is not our benchmark is because we do not have the full variance-covariance matrix as moments M8 and M9 are taken from other papers (based on good identification strategies, e.g. to identify spillovers). We hence assume the off-diagonal terms are zero. Table S.IV shows the match for the targeted moments and Table S.III the estimated parameter values. The results are very similar to our benchmark ones.

In the remaining figures we change the externally calibrated parameters. In all these cases, it is important bearing in mind that wedges represent the gap between what firms would do in the laissez-faire and what the planner induces them to do in the optimal mechanism. Variations in any of these parameters not only change the optimal allocation, they also change what firms would optimally do in the laissez-faire, often in the same direction. As a result, the wedges may not change that much from a change in these parameters; however, the allocations induced could be very different. This is why we show all the wedges and the allocations for each set of parameter values. In addition, total revenues raised by the government and consumer welfare would also be very different since they depend not just on the total innovation produced, but also on the share that goes to consumers.

In Figures S9 to S16, we explore the role of the stochastic type process assumed, some of which was already discussed in the main text. More precisely, Figures S9 and S10 show the wedges for a first-order autoregressive process; Figure S11 and S12 an increasing persistence over the life cycle; Figures S13 to S16 respectively have $p = 0.5$ and $p = 0.9$. The persistence of this stochastic process affects the rate of decay of the wedges very significantly, but not the qualitative findings described above. In addition, a more persistent process increases the ability of the planner to provide dynamic incentives and improves the allocations: there are higher levels of effort and R&D investment for firms of all productivities.

Figures S17 to S20 show the changes induced by higher or lower values of β. Higher β represents a higher degree of market power, as it increases the markup over marginal costs that the intermediate good producer can charge. At the same time, it also means that the quality of each differentiated product is valued more by consumers. On balance, there is more investment in R&D and more effort at the optimum when β is higher.

Figures S21 and S24 consider higher rates of depreciation of innovation, of $\delta = 0.15$ and $\delta = 0.3$ respectively. The higher the rate of depreciation, the higher the wedges have to be to induce firms to invest sufficiently much (relative to what they would do if left to choose). Naturally, the higher the rate at which knowledge depreciates and the lower the optimal investments, step sizes, and resulting innovation that can be stimulated.

Finally, Figures S27 and S28 show what happens when the cost of R&D is less convex, i.e., when $\eta = 1$. This barely changes the wedges, as they represent the share of costs that is subsidized. However, as expected, the level of R&D effort and incentives that can be incentivized are larger when costs are less convex.

Table S.III: Parameter Values using Two-step GMM

Parameter	Symbol	Value
External Calibration		
Interest rate	R	1.05
Intangibles depreciation	δ	0.1
Knowledge share	β	0.15
R&D cost elasticity	η	1.5
Level of types	μ_θ	0.00
Initial R&D stock	r_0	1.0
Program horizon	T	30
Internal Calibration		
R&D share	α	0.48
R&D-type substitution	$\rho_{\theta r}$	1.84
Type variance	σ_ϵ	0.342
Type persistence	$\tilde{p}$	0.69
Scale of disutility	κ_l	0.72
Scale of R&D cost	κ_r	0.061
Effort cost elasticity	γ	0.94
Support width for θ_1	Θ^1	1.75
Production externality	ζ	0.018

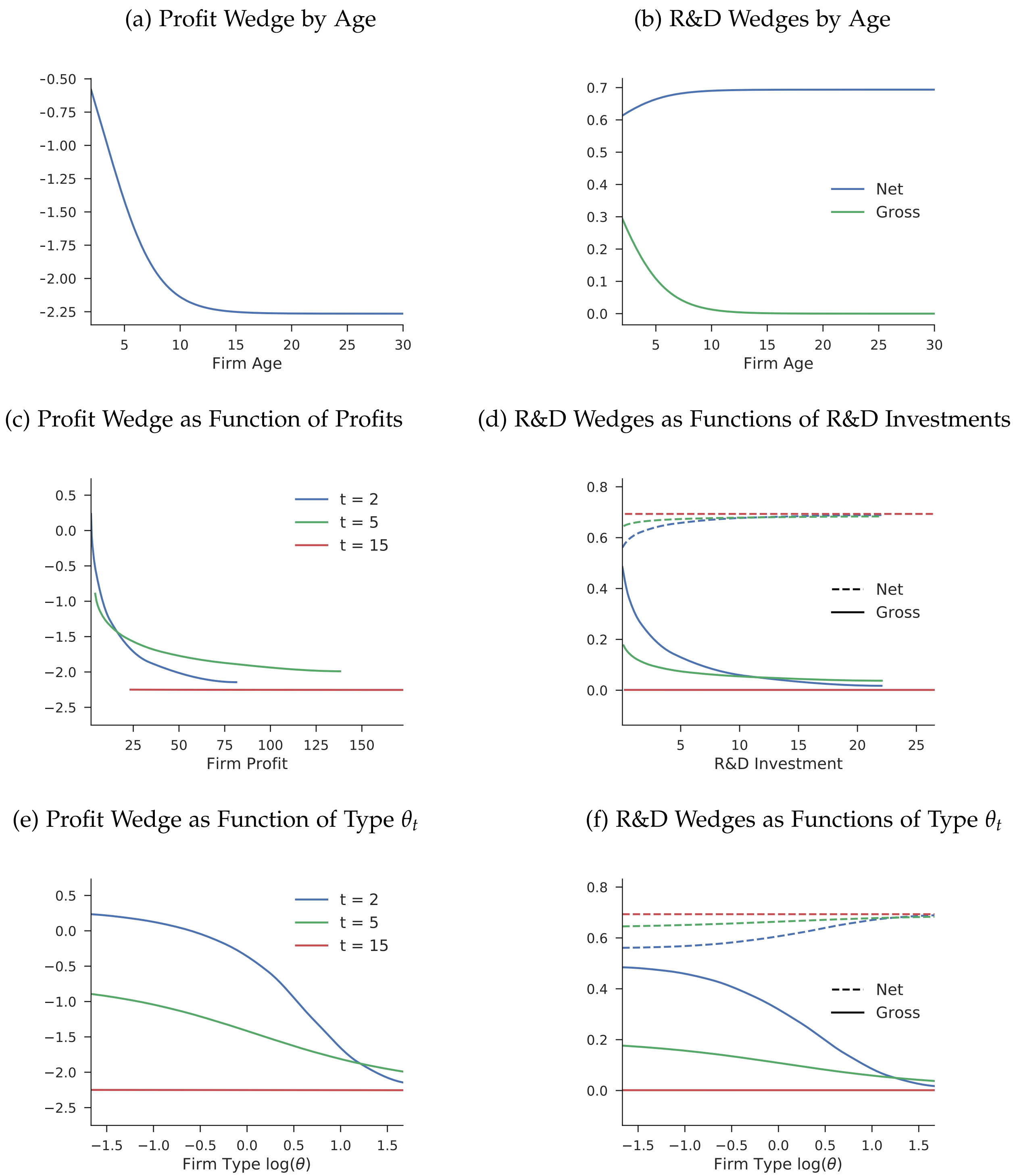
Table S.IV: Moments using Two-step GMM

Moment	Target	Simulation	Standard Error
M1. Patent quality-R&D elasticity	0.88	0.96	(0.0009)
M2. R&D/Sales mean	0.041	0.034	(0.0025)
M3. Sales growth (DHS) mean	0.06	0.07	(0.005)
M4. Within-firm patent quality coeff of var	0.63	0.79	(0.0017)
Across-firm patent quality coeff of var:			
M5. Young firms	1.06	1.04	(0.0012)
M6. Older firms	0.99	0.89	(0.0016)
M7. Patent quality young/old	1.04	1.03	(0.0048)
M8. Spillover coefficient	0.191	0.190	(0.046)
M9. Elasticity of R&D investment to cost	- 0.35	-0.34	(0.101)

References

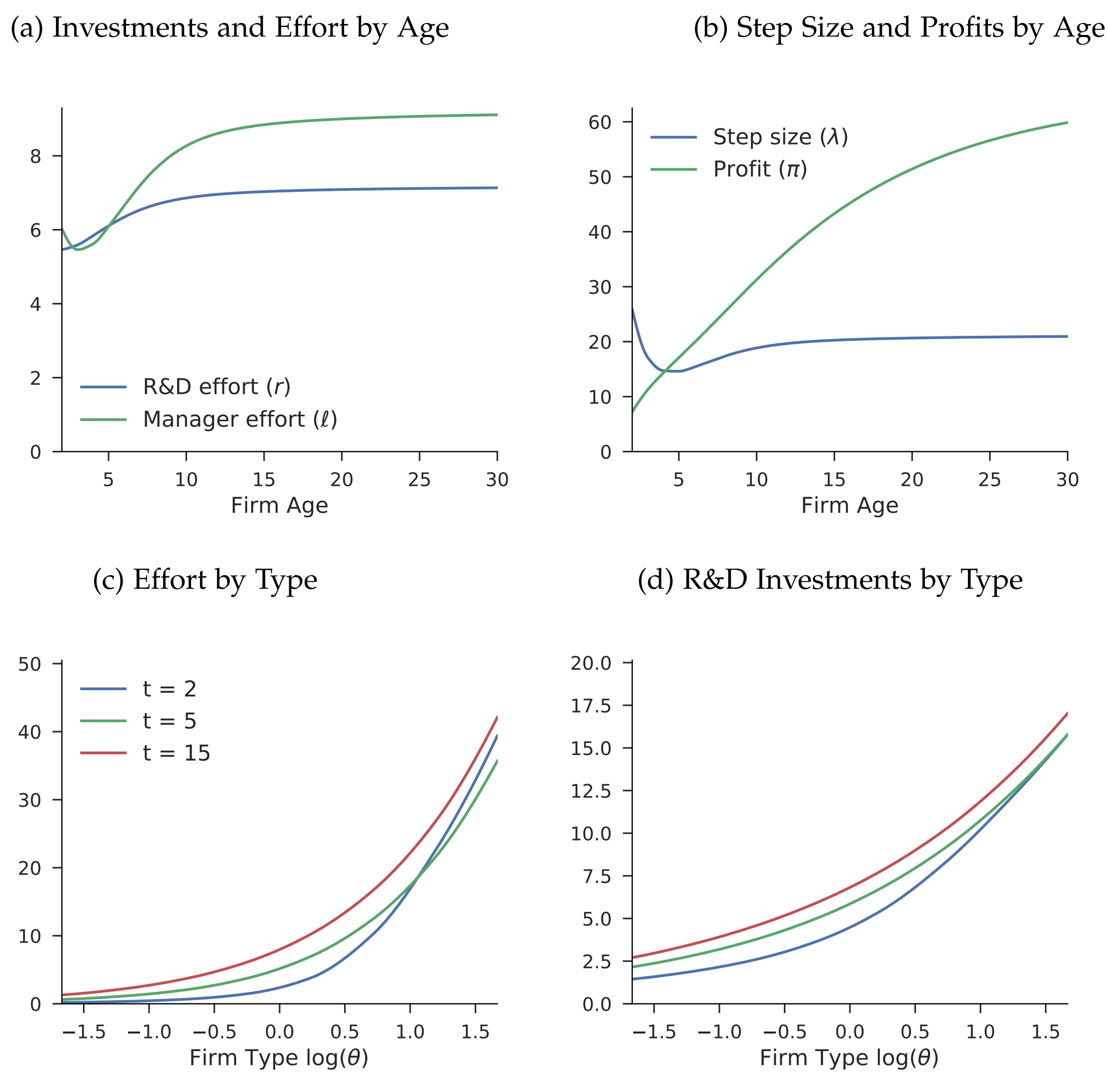
Bloom, Nicholas, Mark Schankerman, and John Van Reenen (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica 81*(4), 1347–1393.

Figure S1: Optimal Profit and R&D Wedges with Quantity Control



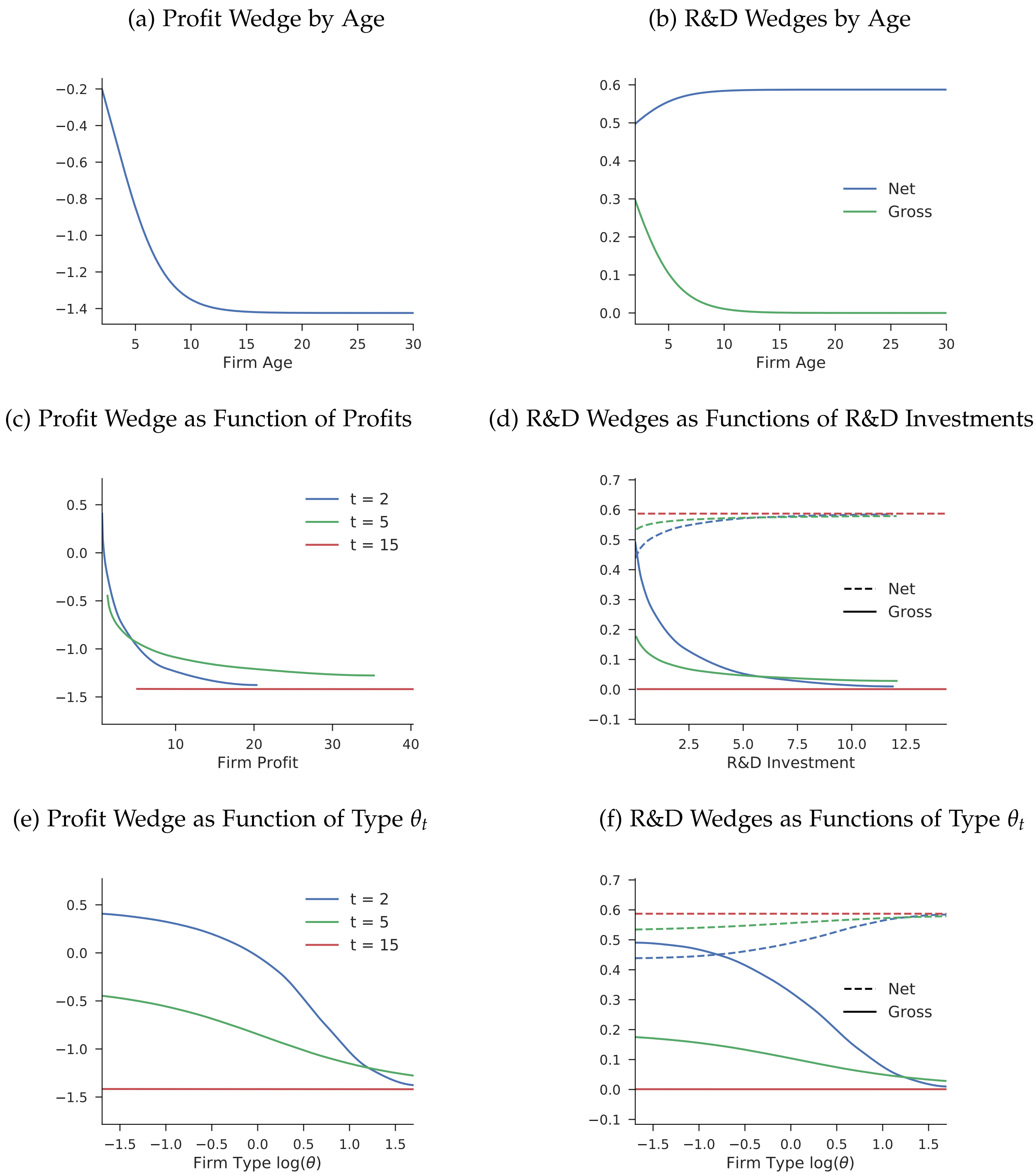
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S2: Optimal Allocations with Quantity Control



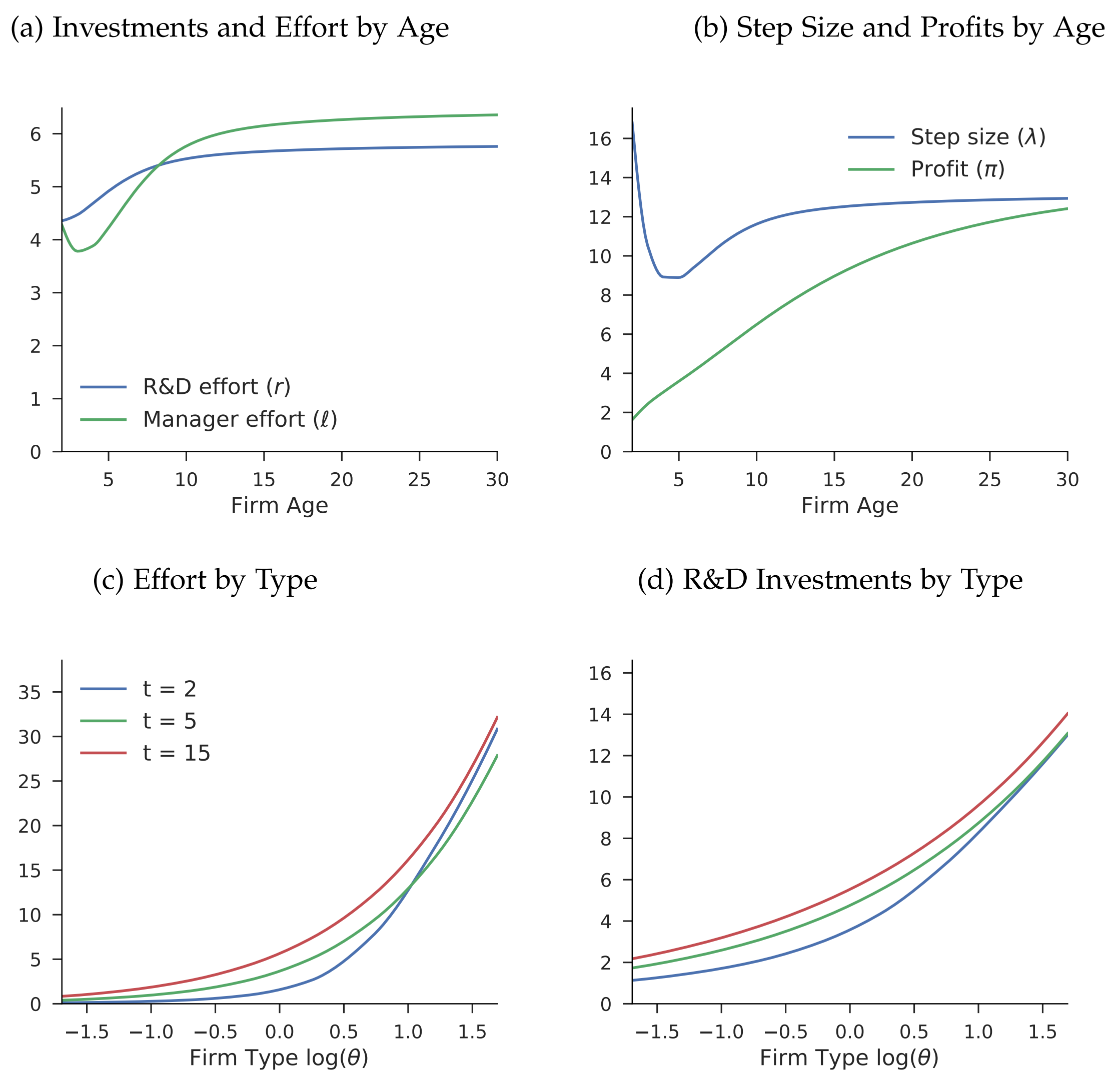
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S3: Optimal Profit and R&D Wedges for COMPUSTAT Publicly Traded Firms



Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

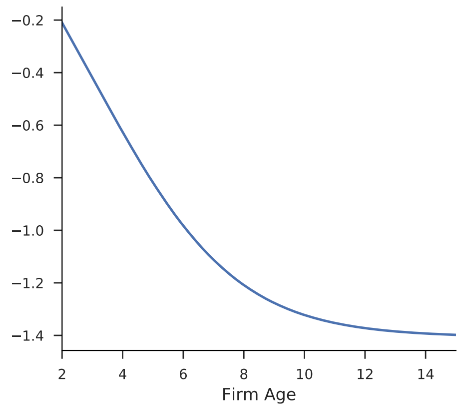
Figure S4: Optimal Allocations for COMPUSTAT Publicly Traded Firms



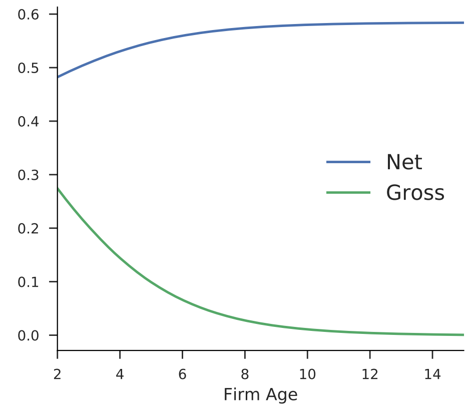
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S5: Optimal Profit and R&D Wedges for Finite Firm Life Cycle $T = 15$

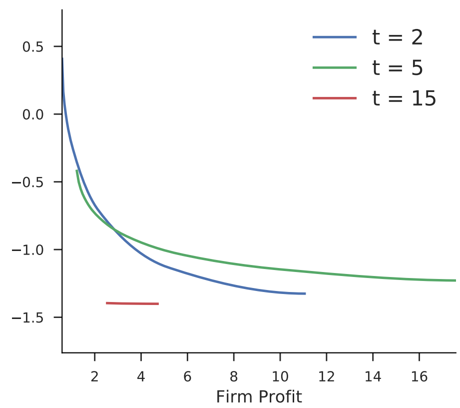
(a) Profit Wedge by Age



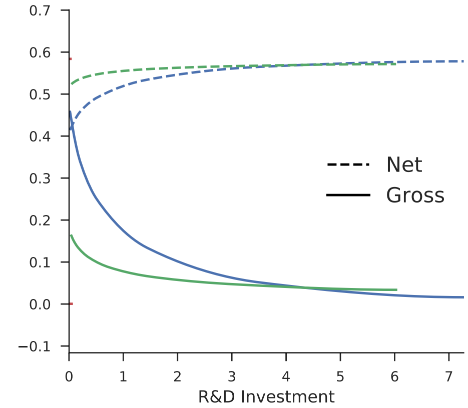
(b) R&D Wedges by Age



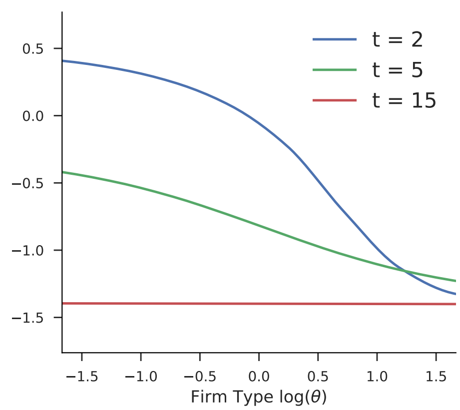
(c) Profit Wedge as Function of Profits



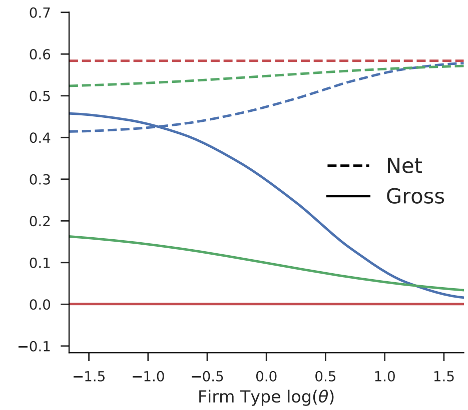
(d) R&D Wedges as Functions of R&D Investments



(e) Profit Wedge as Function of Type θ_t

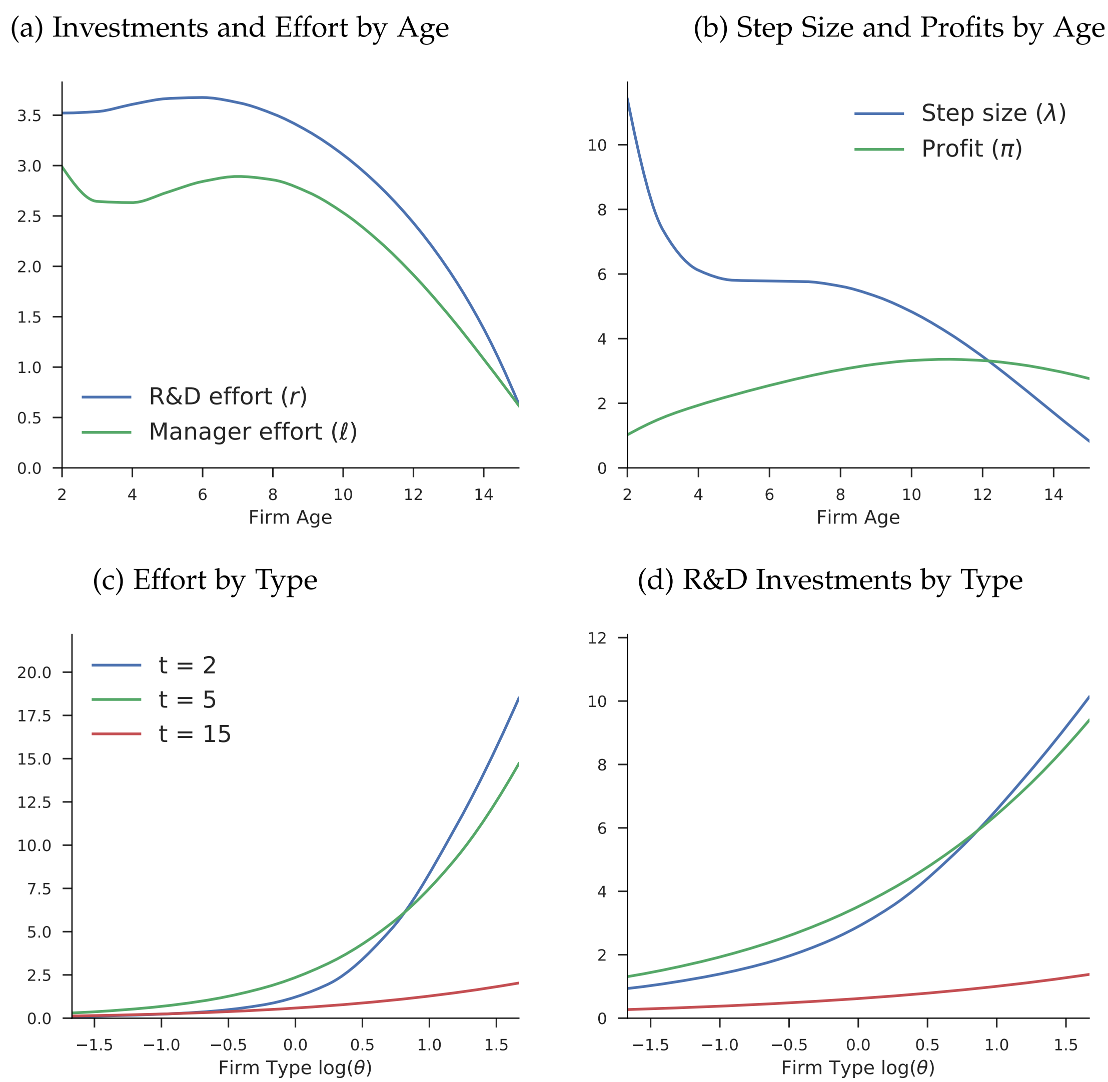


(f) R&D Wedges as Functions of Type θ_t



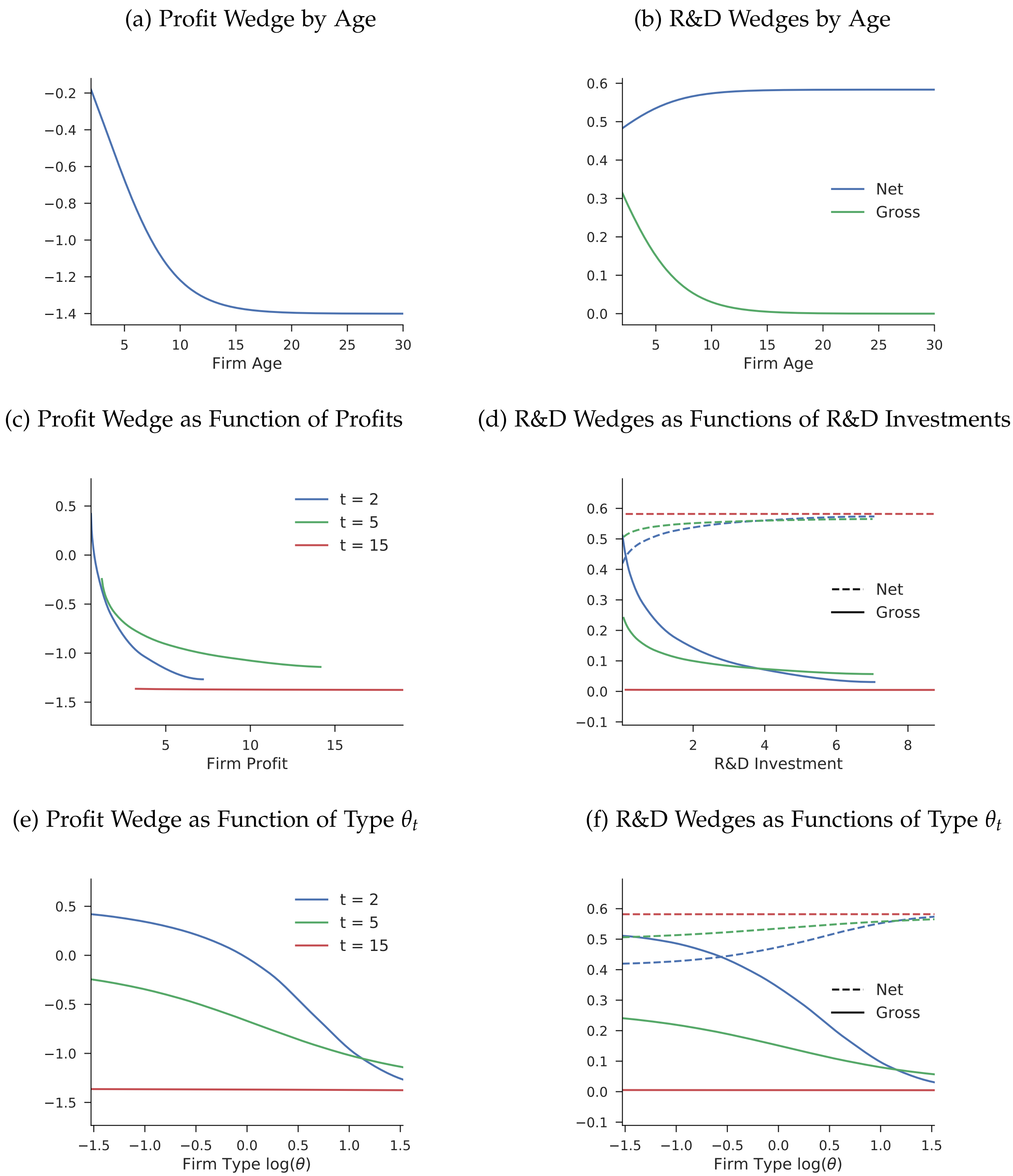
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S6: Optimal Allocations for Finite Firm Life Cycle $T = 15$



Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

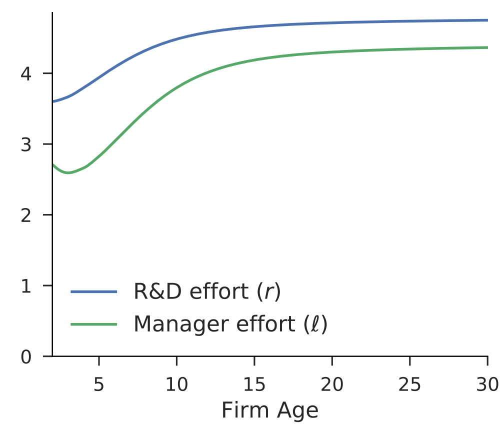
Figure S7: Optimal Profit and R&D Wedges using Two-step GMM



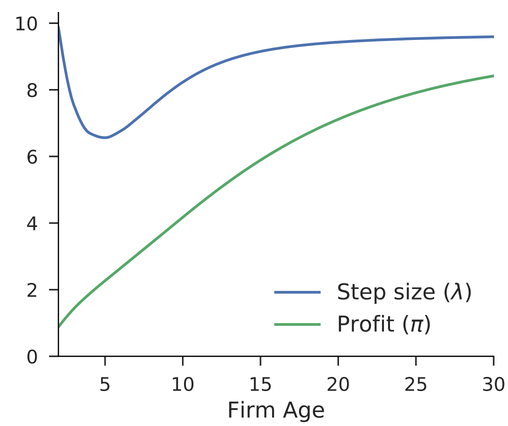
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S8: Optimal Allocations using Two-step GMM

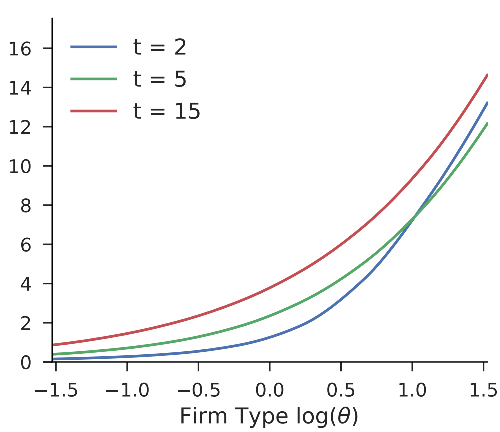
(a) Investments and Effort by Age



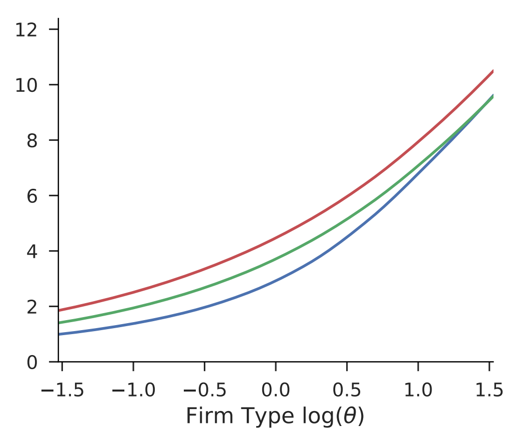
(b) Step Size and Profits by Age



(c) Effort by Type



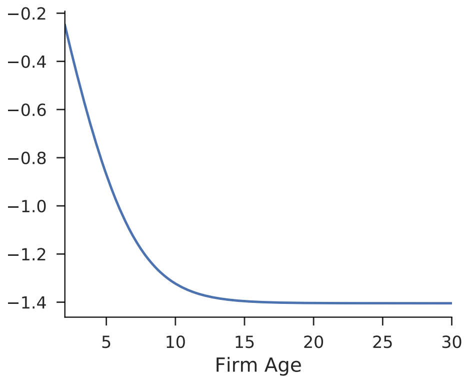
(d) R&D Investments by Type



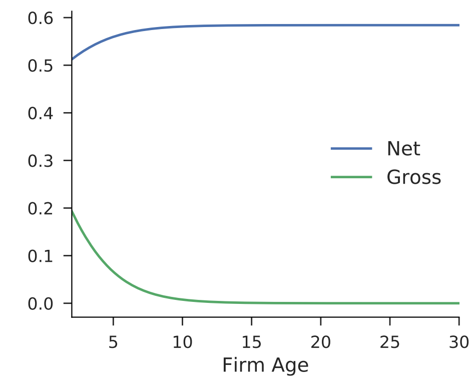
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S9: Optimal Profit and R&D Wedges with an Autoregressive Process

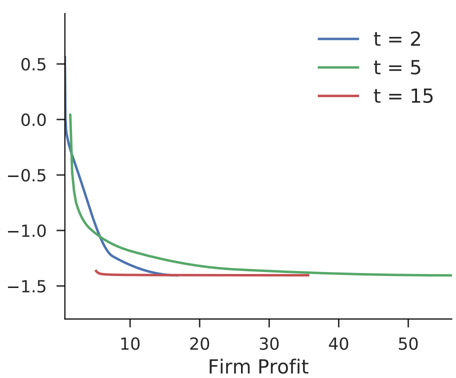
(a) Profit Wedge by Age



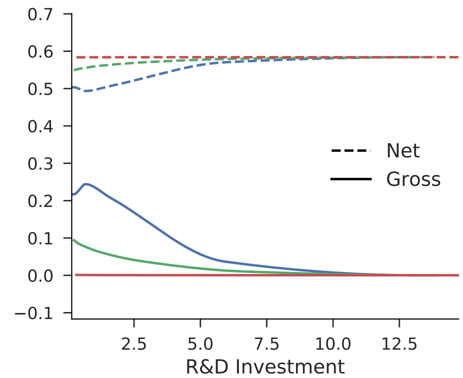
(b) R&D Wedges by Age



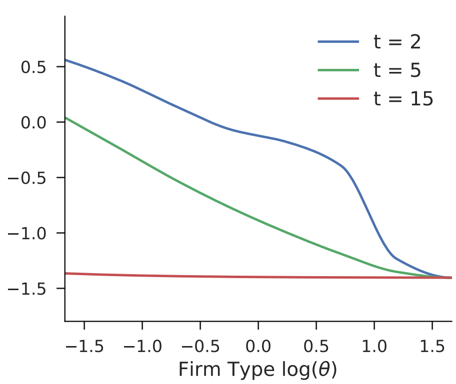
(c) Profit Wedge as Function of Profits



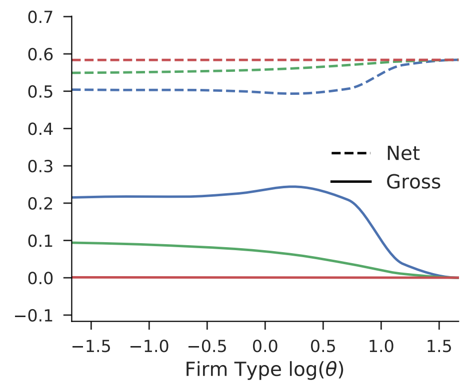
(d) R&D Wedges as Functions of R&D Investments



(e) Profit Wedge as Function of Type θ_t

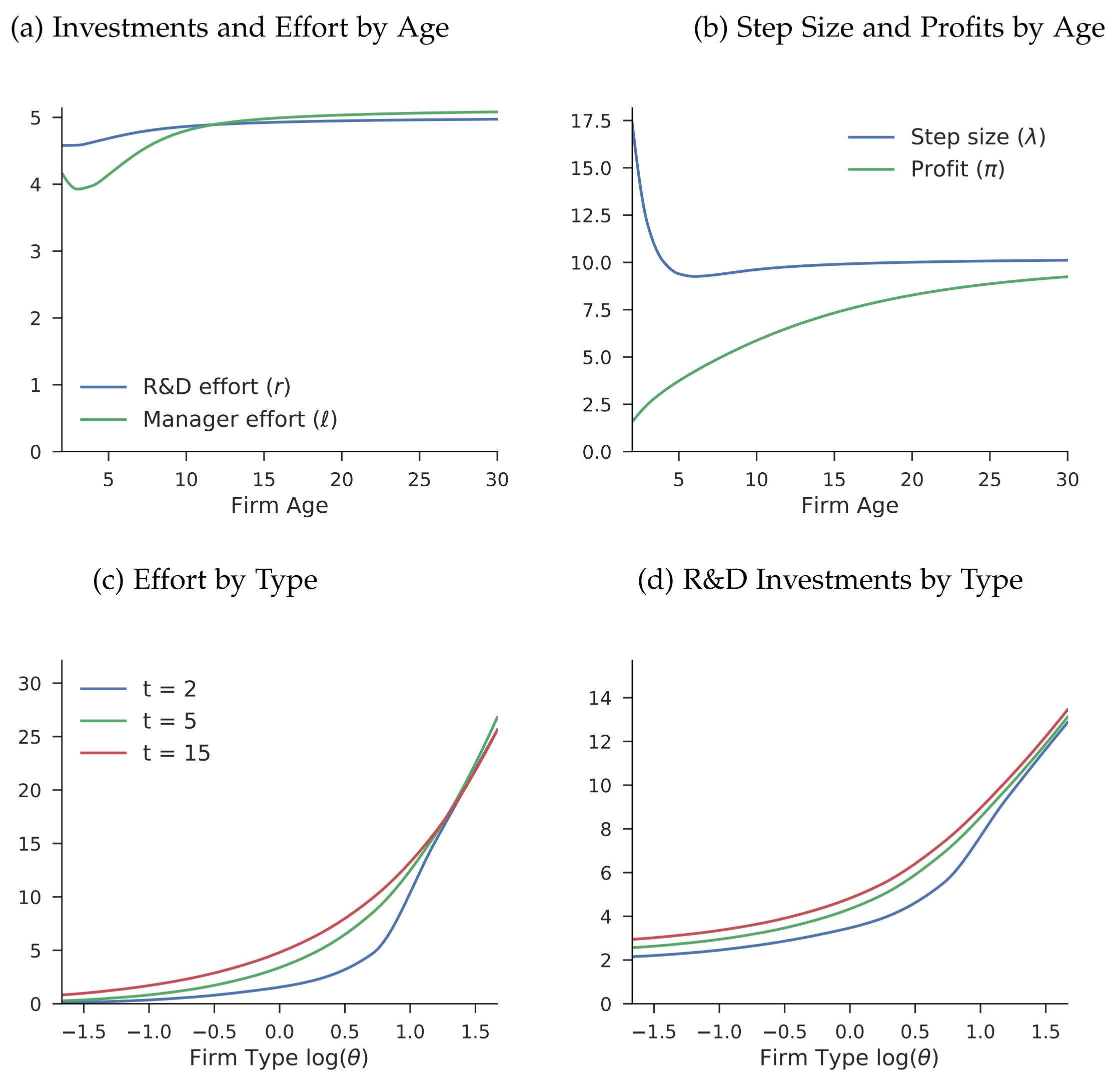


(f) R&D Wedges as Functions of Type θ_t



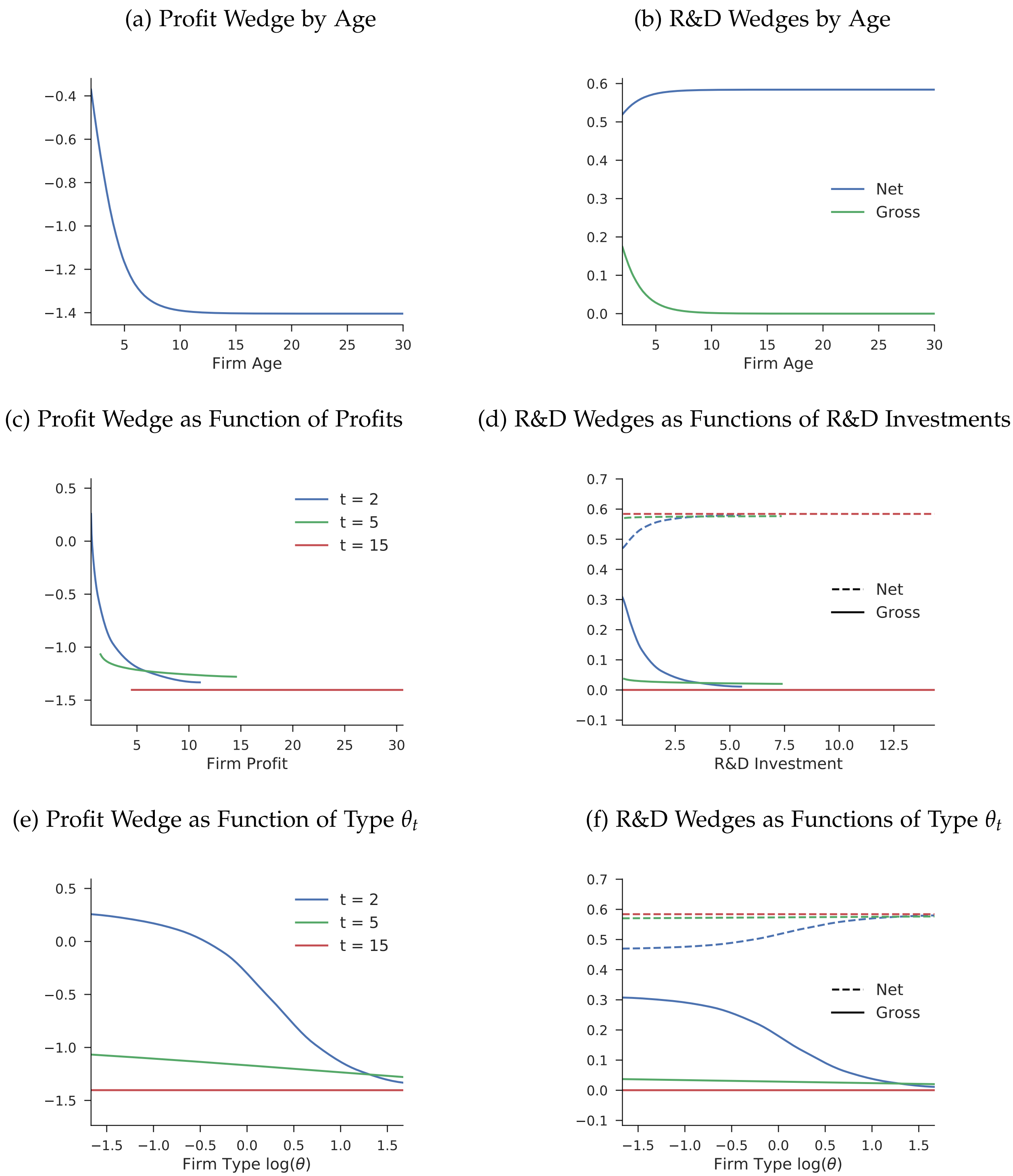
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S10: Optimal Allocations with an Autoregressive Process



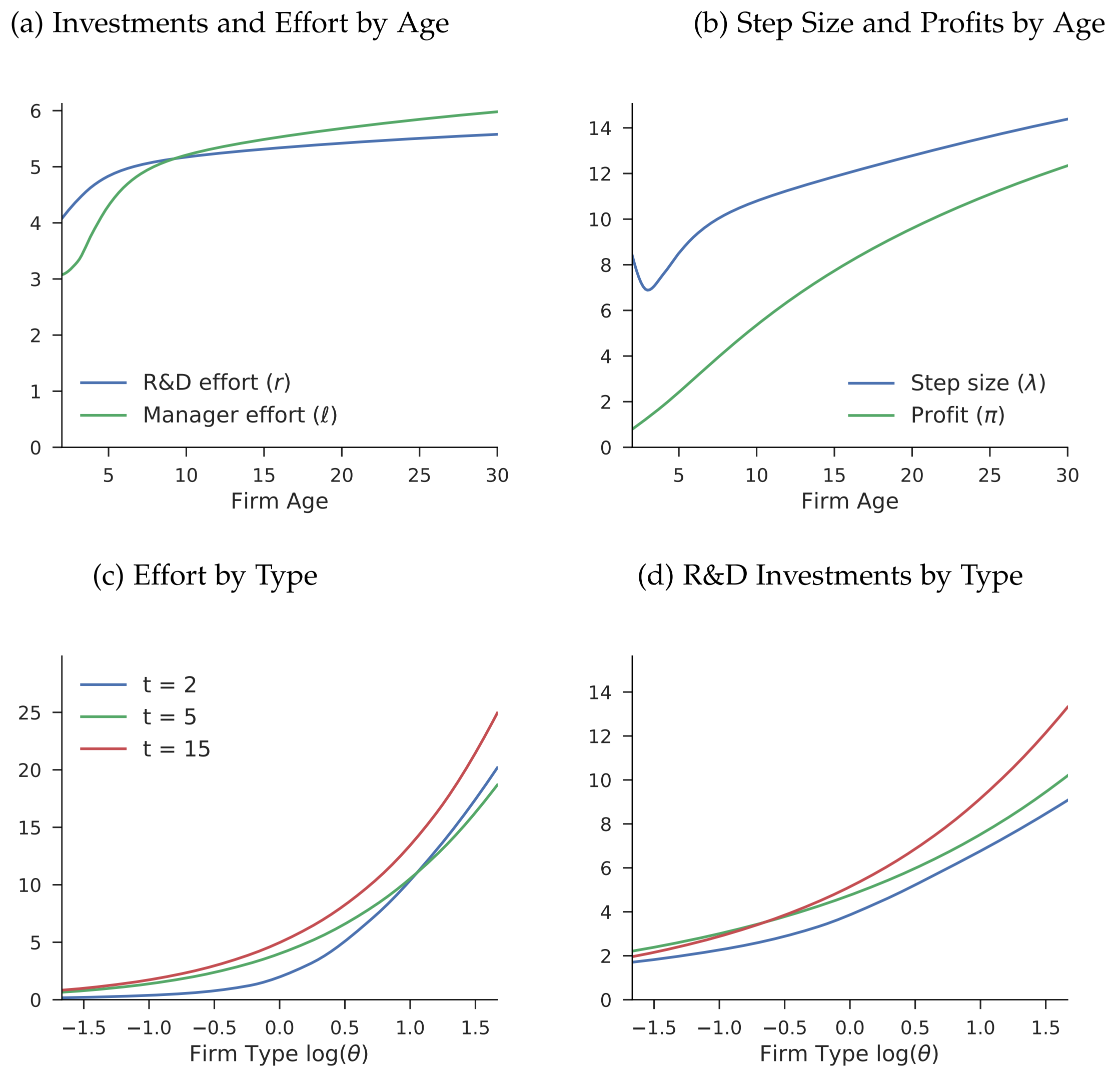
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S11: Optimal Profit and R&D Wedges with Increasing Persistence p



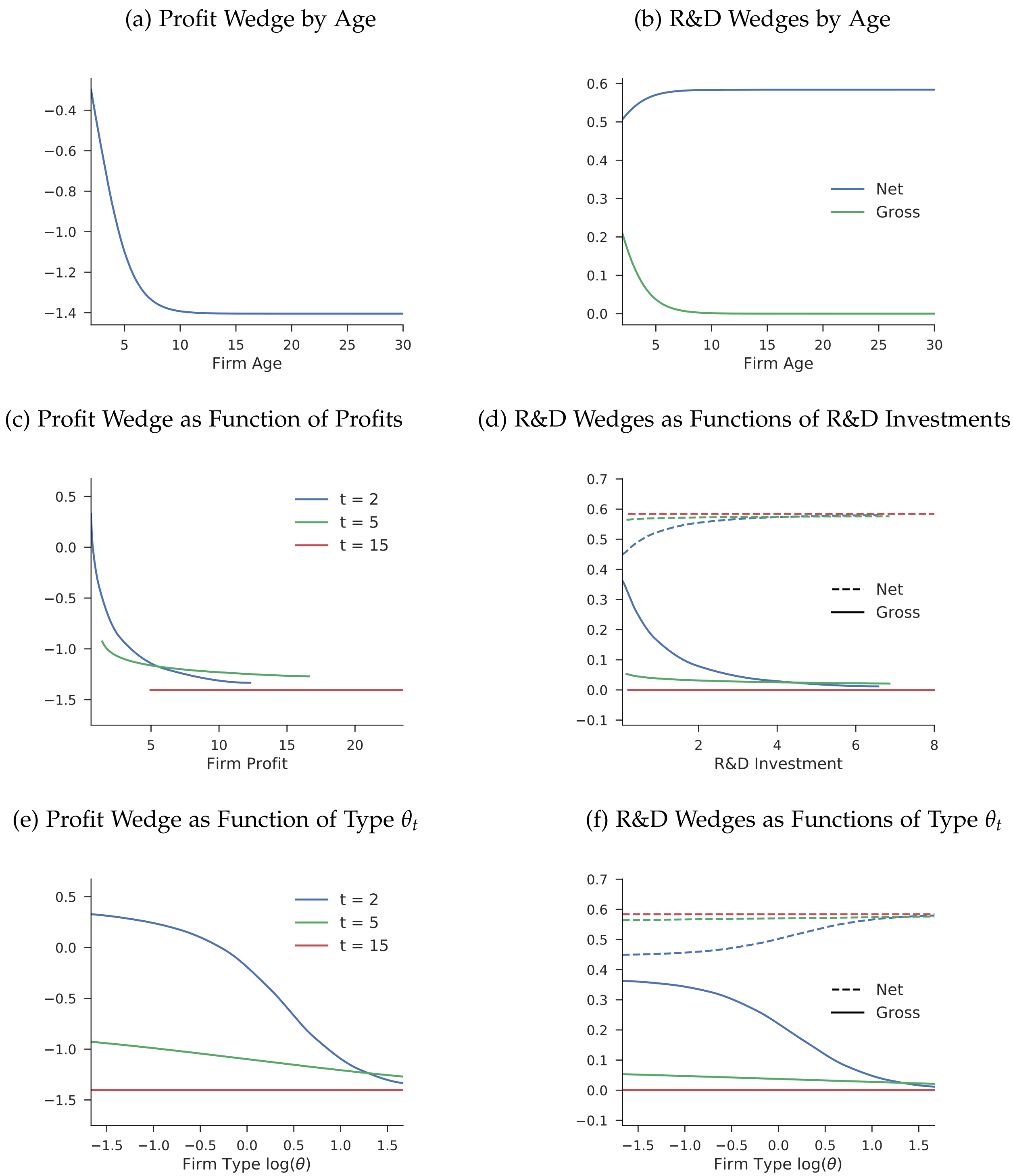
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S12: Optimal Allocations with Increasing Persistence p



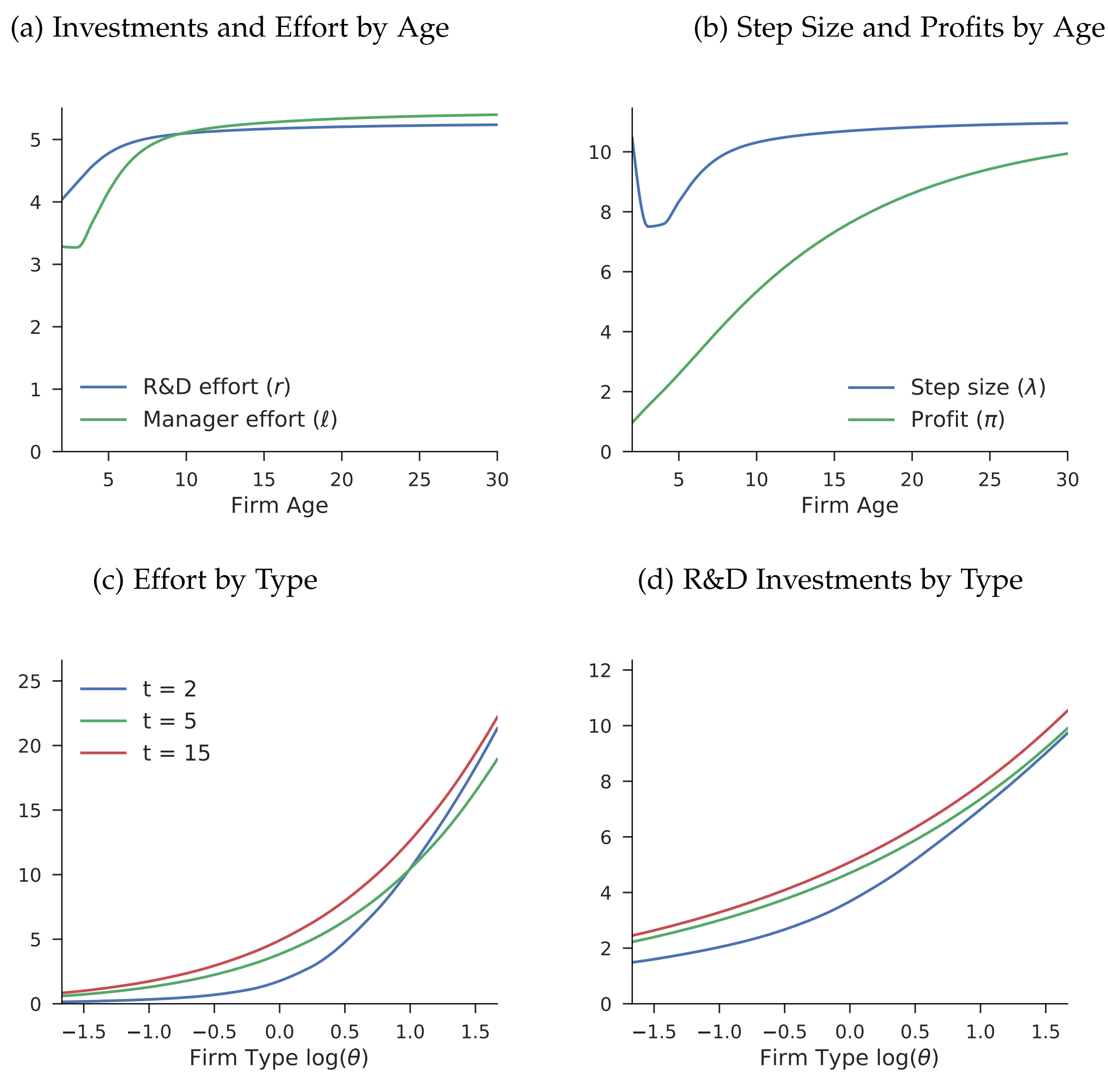
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S13: Optimal Profit and R&D Wedges with Persistence $p = 0.5$



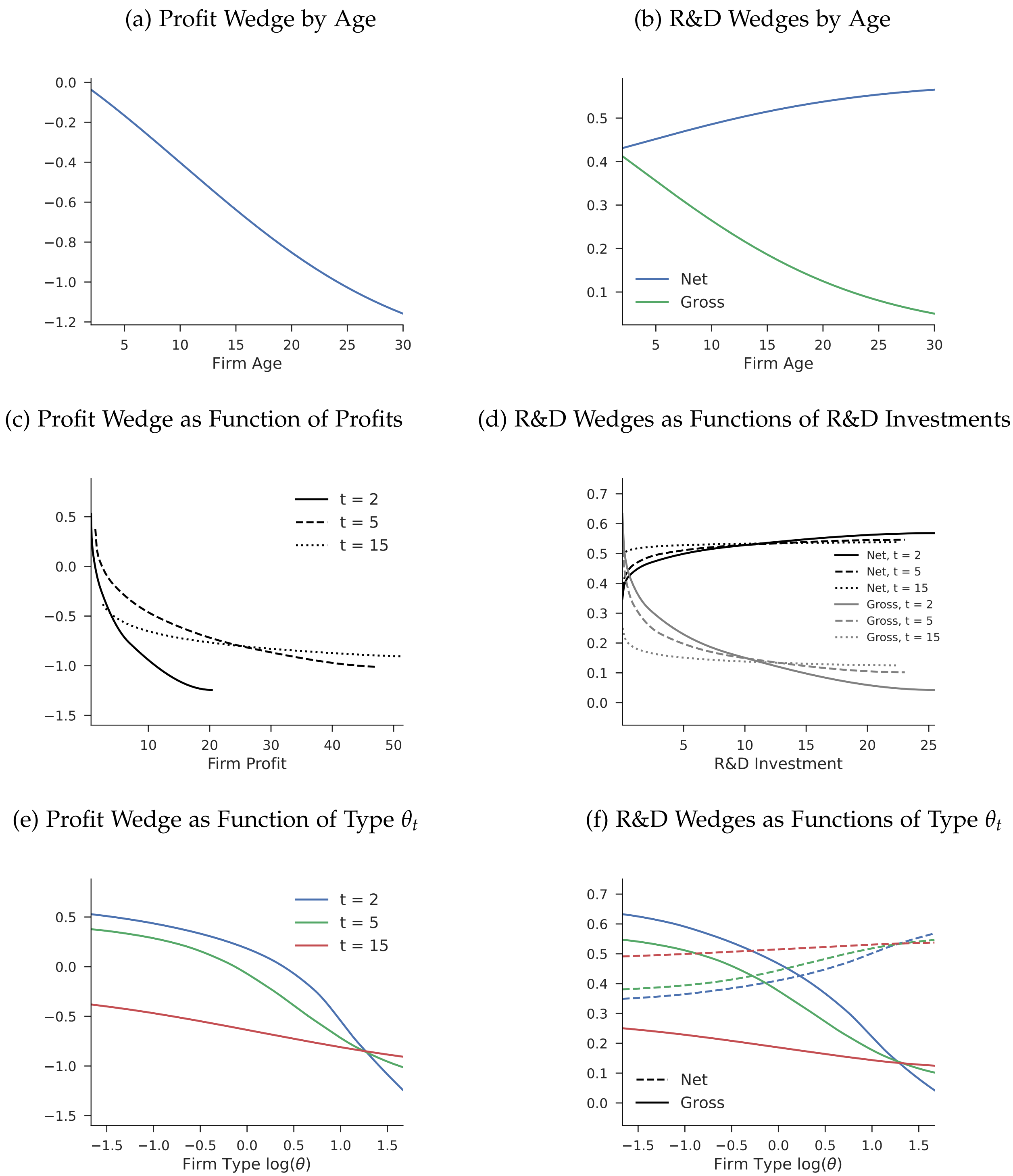
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S14: Optimal Allocations with Persistence $p = 0.5$



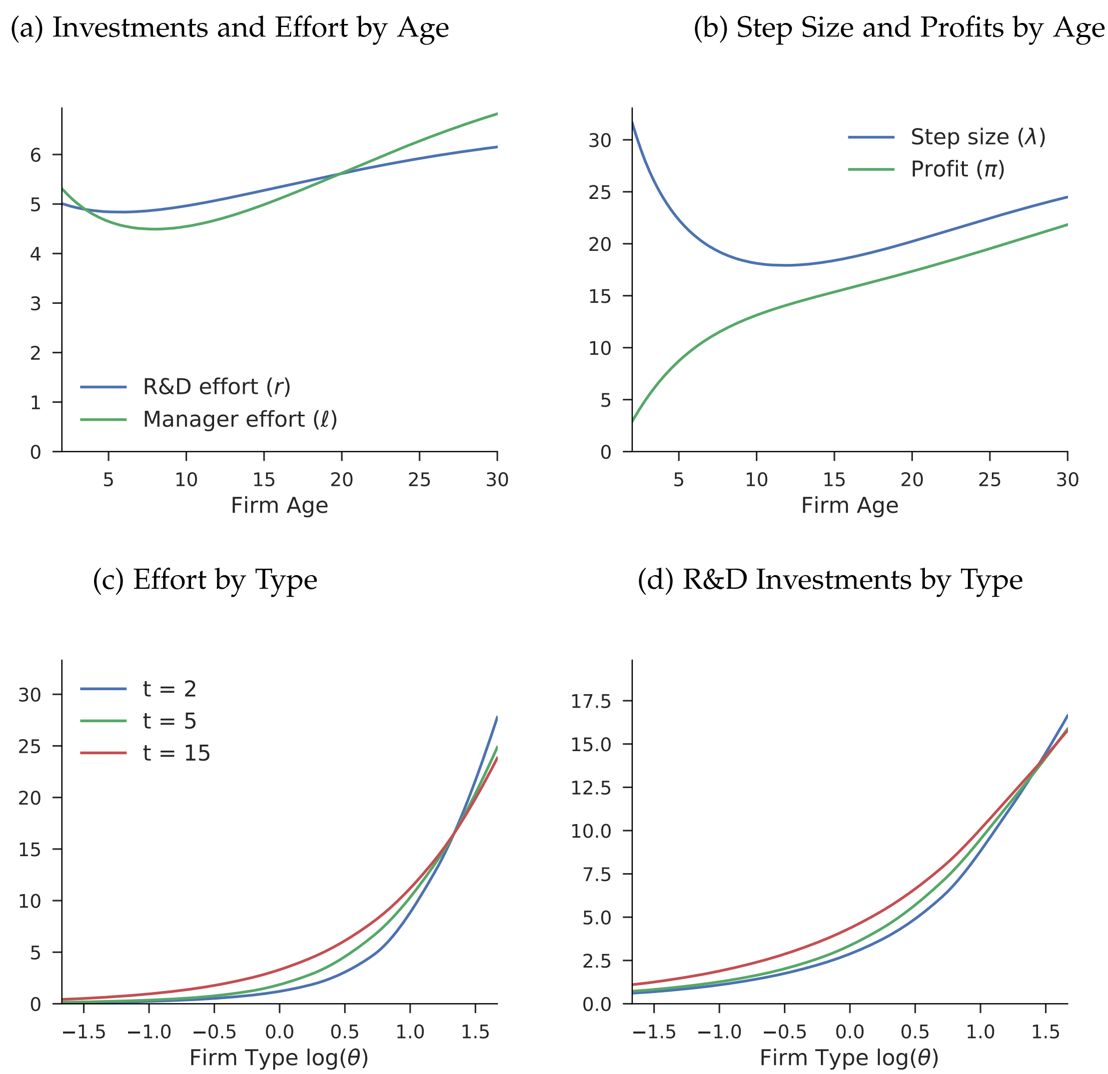
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S15: Optimal Profit and R&D Wedges with Persistence $p = 0.9$



Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

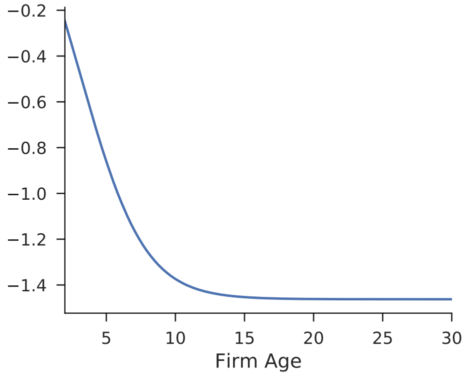
Figure S16: Optimal Allocations with Persistence $p = 0.9$



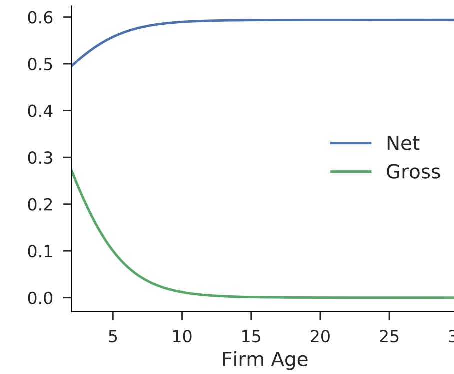
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S17: Optimal Profit and R&D Wedges for $\beta = 0.10$

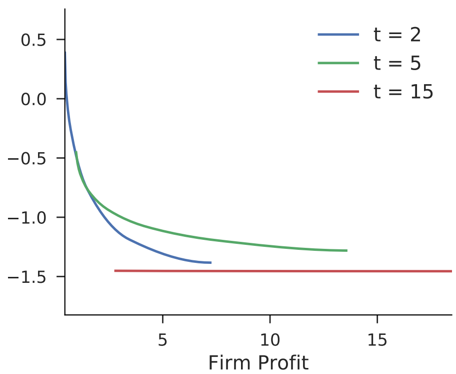
(a) Profit Wedge by Age



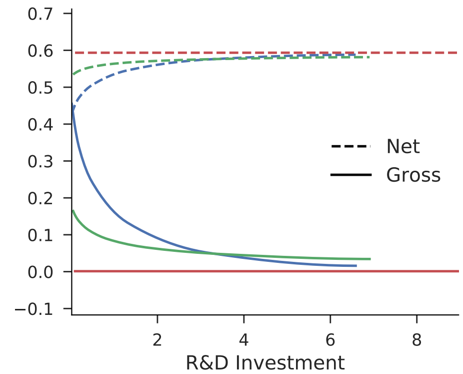
(b) R&D Wedges by Age



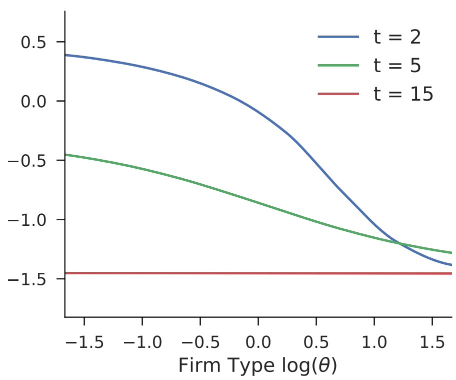
(c) Profit Wedge as Function of Profits



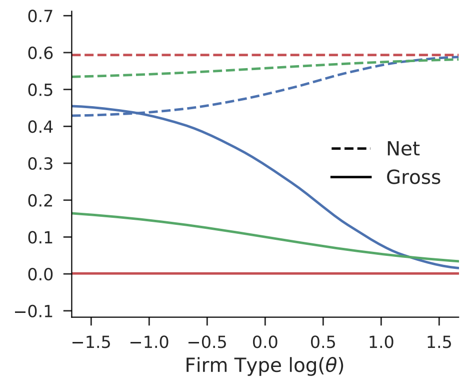
(d) R&D Wedges as Functions of R&D Investments



(e) Profit Wedge as Function of Type θ_t



(f) R&D Wedges as Functions of Type θ_t



Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

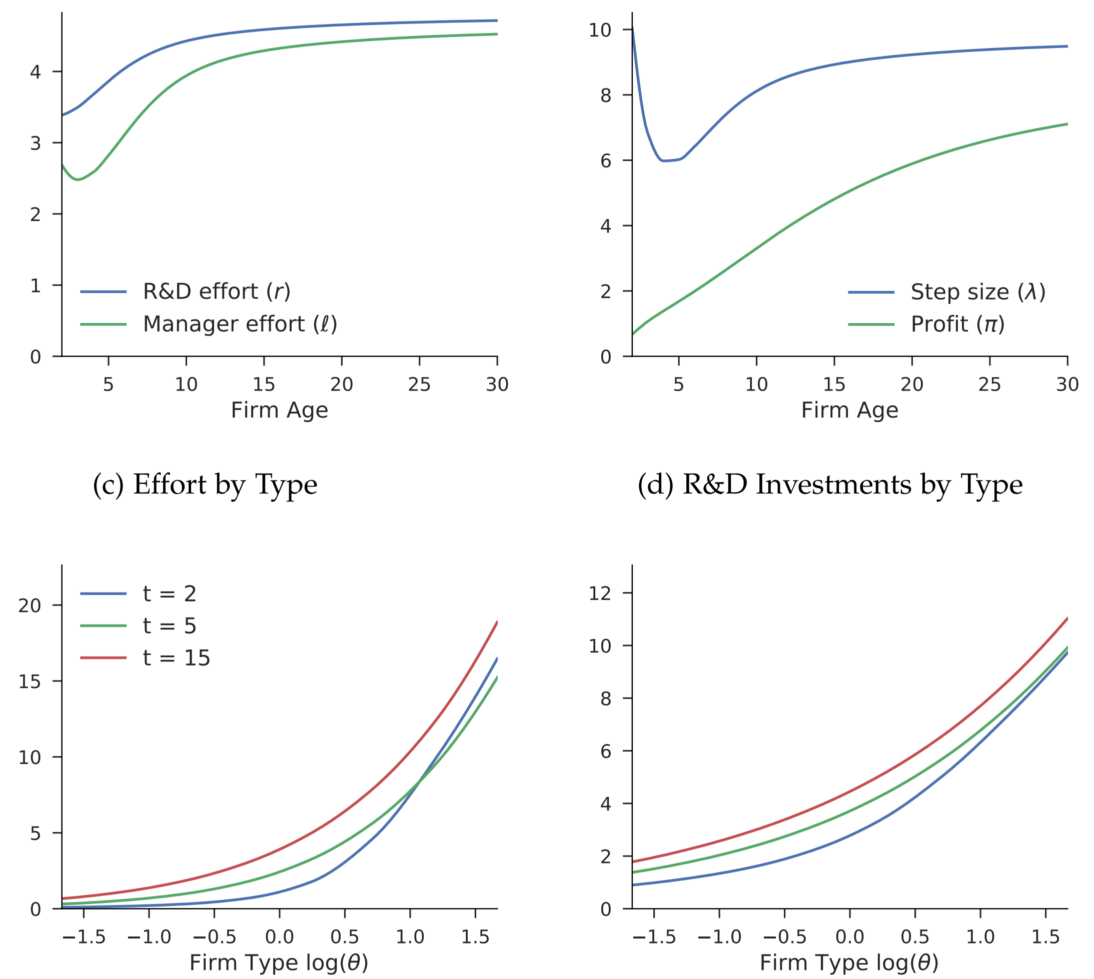
FIGURE S18: OPTIMAL ALLOCATIONS FOR $\beta = 0.10$

(a) Investments and Effort by Age

(b) Step Size and Profits by Age

(c) Effort by Type

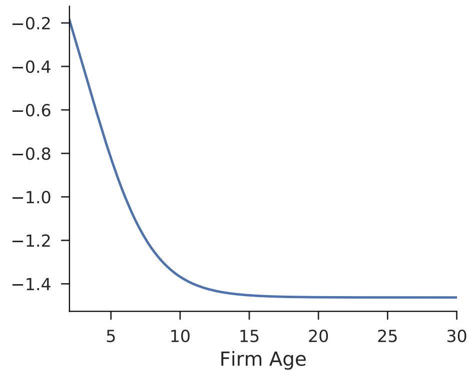
(d) R&D Investments by Type



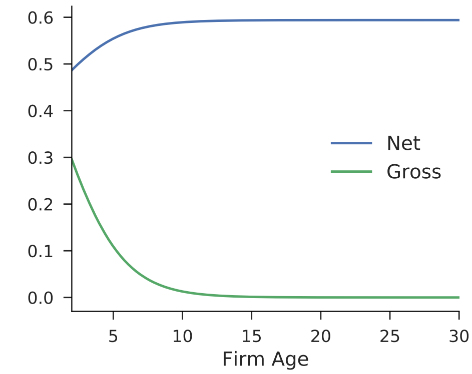
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

FIGURE S19: OPTIMAL PROFIT AND R&D WEDGES FOR $\beta = 0.25$

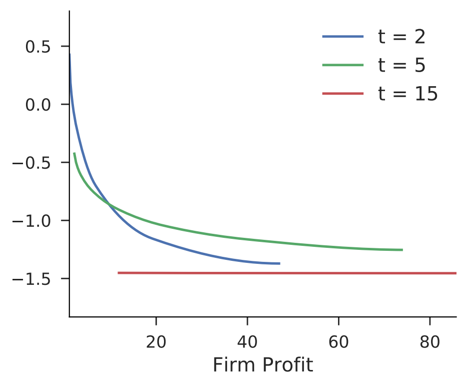
(a) Profit Wedge by Age



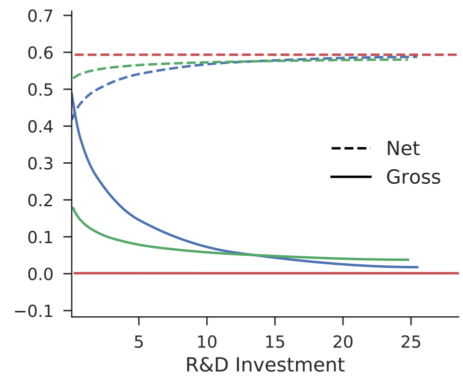
(b) R&D Wedges by Age



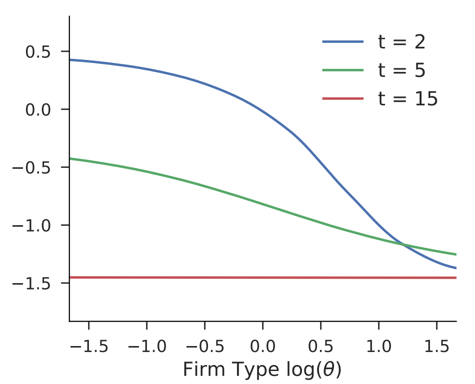
(c) Profit Wedge as Function of Profits



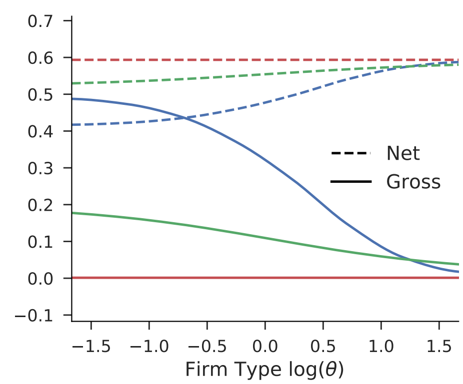
(d) R&D Wedges as Functions of R&D Investments



(e) Profit Wedge as Function of Type θ_t

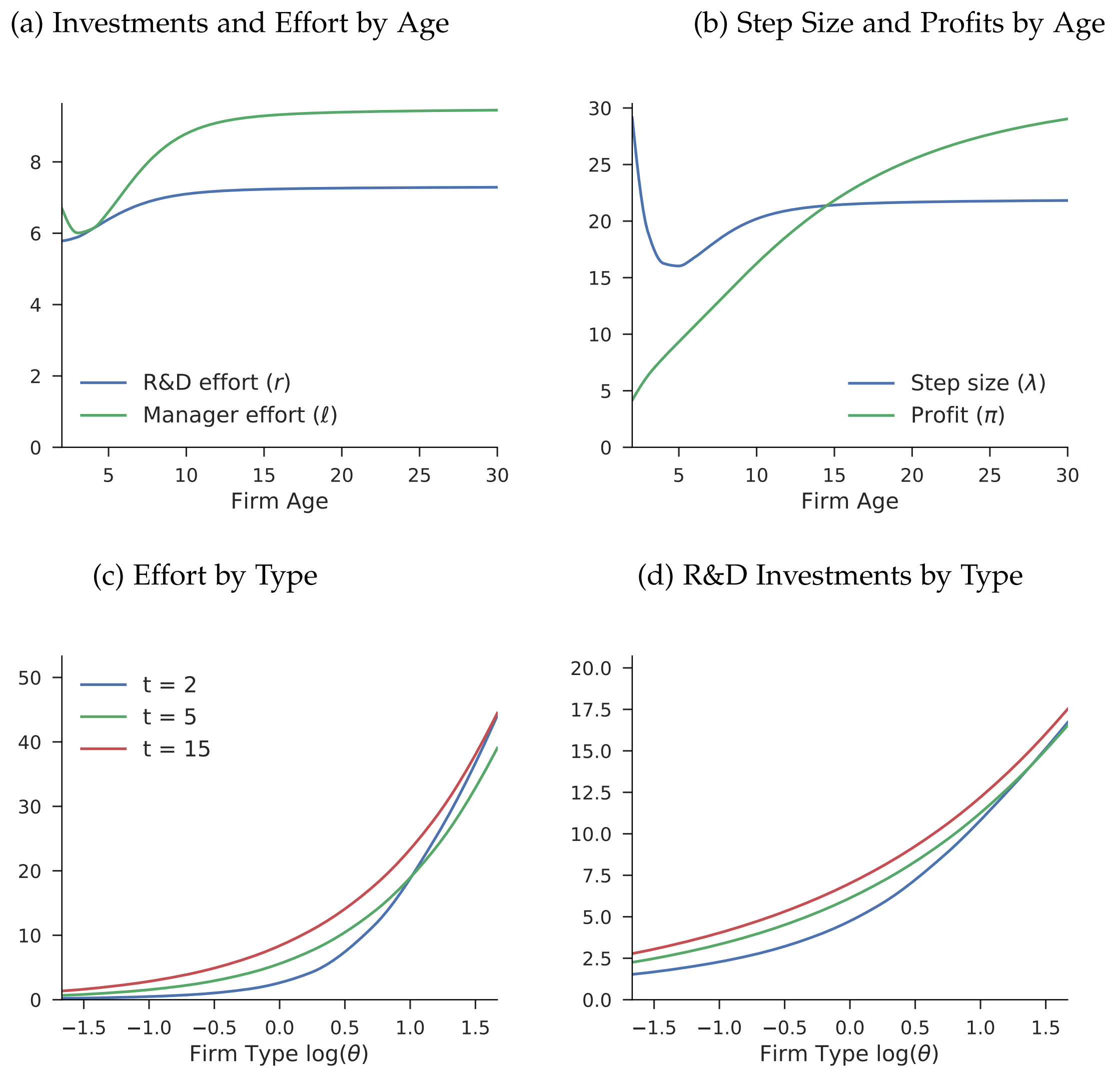


(f) R&D Wedges as Functions of Type θ_t



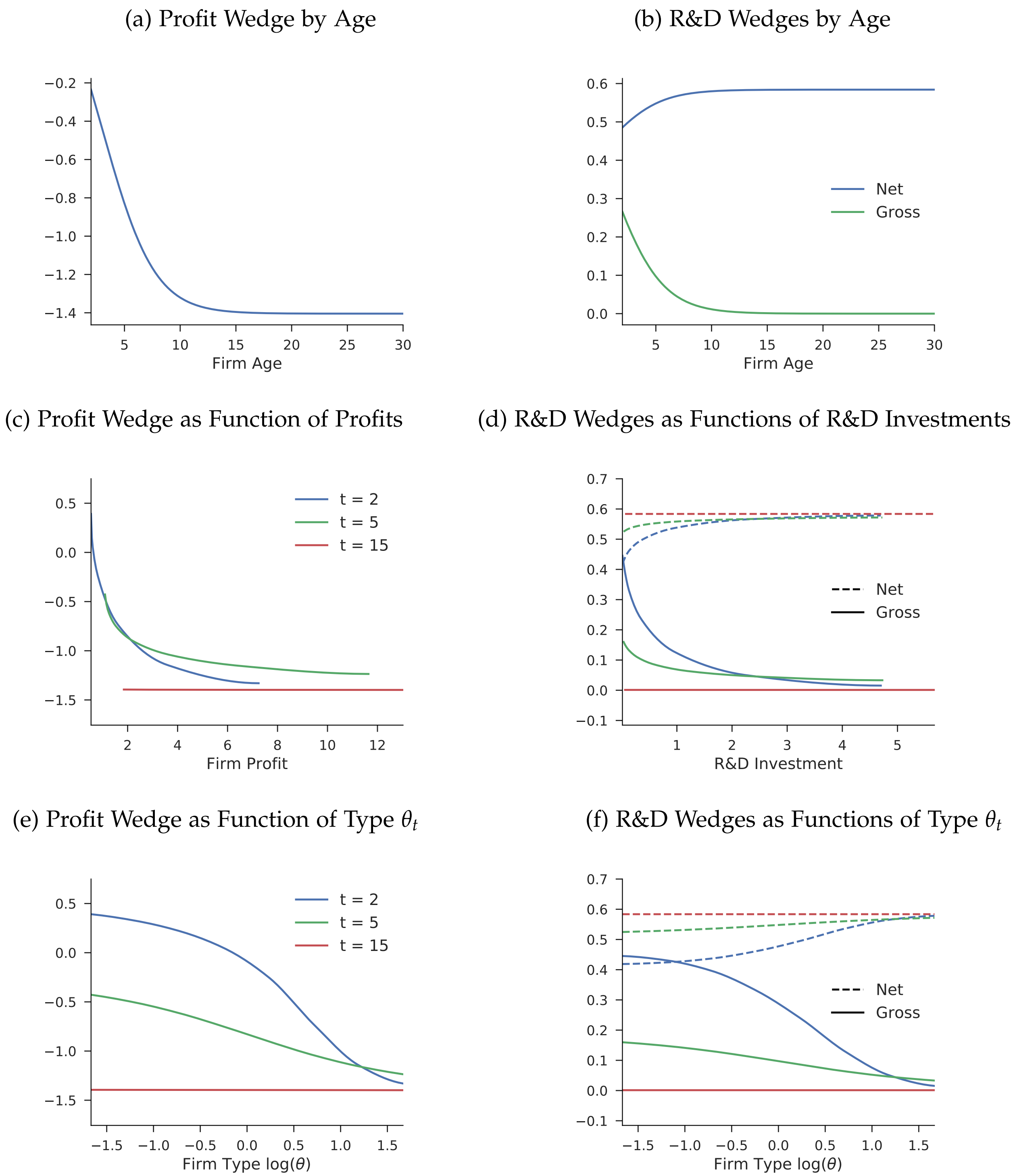
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S20: Optimal Allocations for $\beta = 0.25$



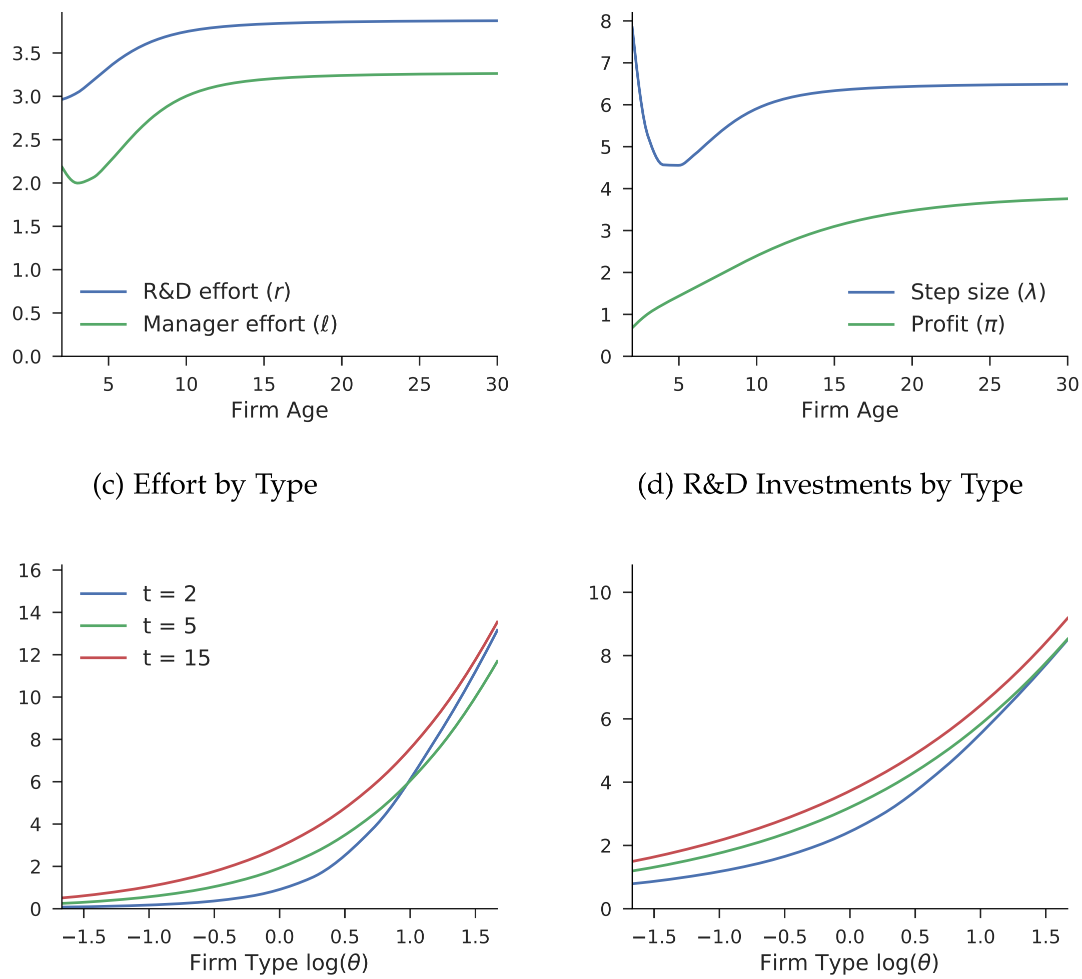
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S21: Optimal Profit and R&D Wedges for $\delta = 0.15$



Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S22: Optimal Allocations for $\delta = 0.15$



(a) Investments and Effort by Age

(b) Step Size and Profits by Age

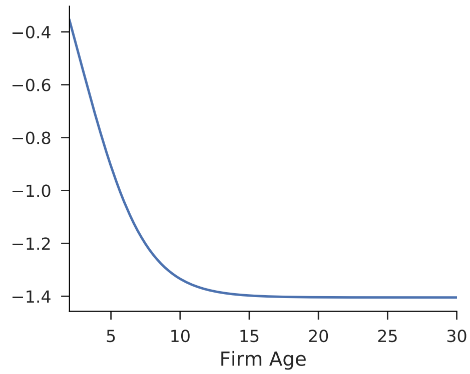
(c) Effort by Type

(d) R&D Investments by Type

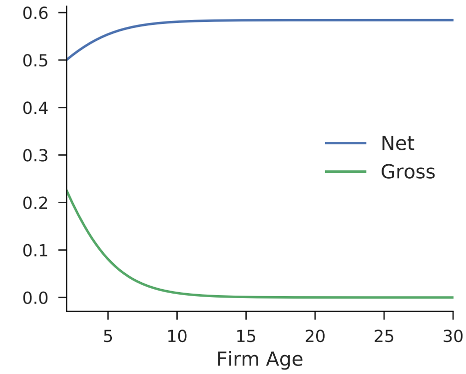
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S23: Optimal Profit and R&D Wedges for $\delta = 0.3$

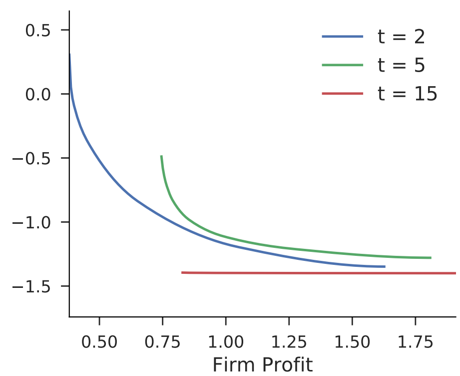
(a) Profit Wedge by Age



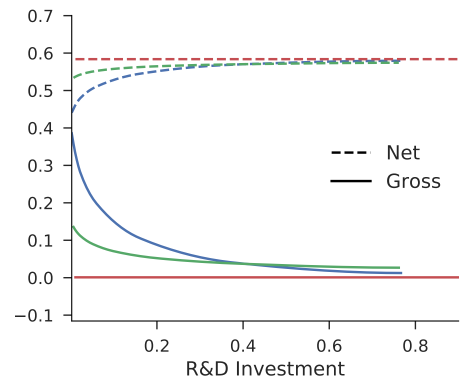
(b) R&D Wedges by Age



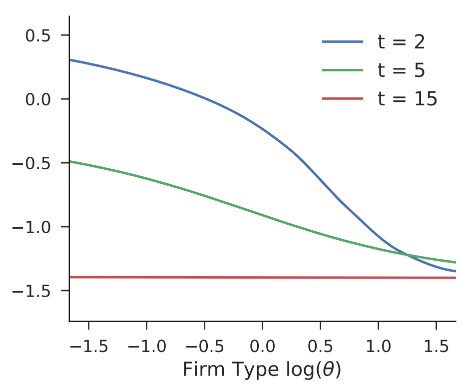
(c) Profit Wedge as Function of Profits



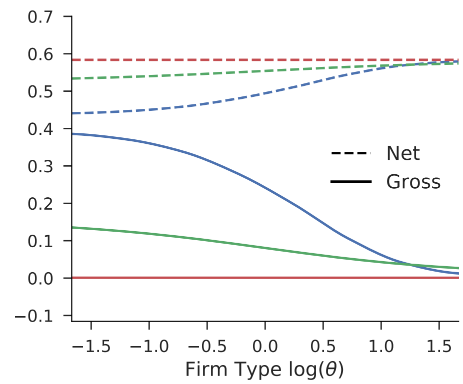
(d) R&D Wedges as Functions of R&D Investments



(e) Profit Wedge as Function of Type θ_t



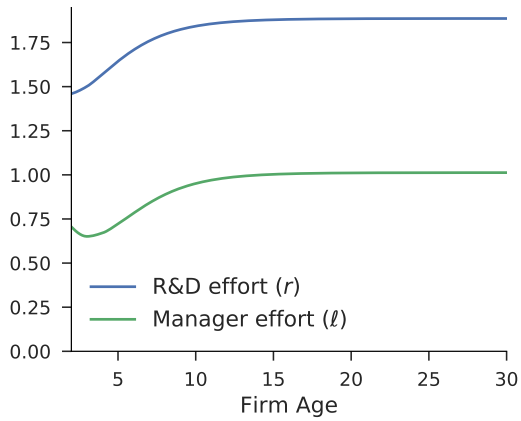
(f) R&D Wedges as Functions of Type θ_t



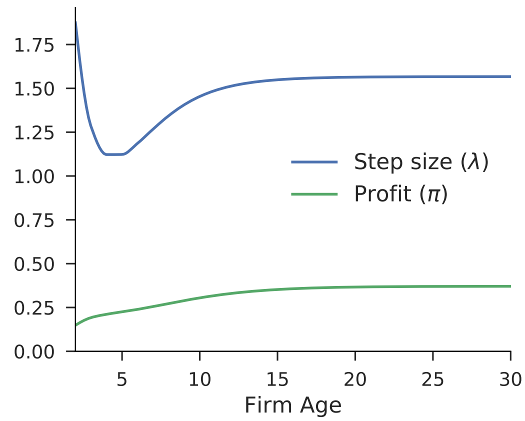
Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S24: Optimal Allocations for $\delta = 0.3$

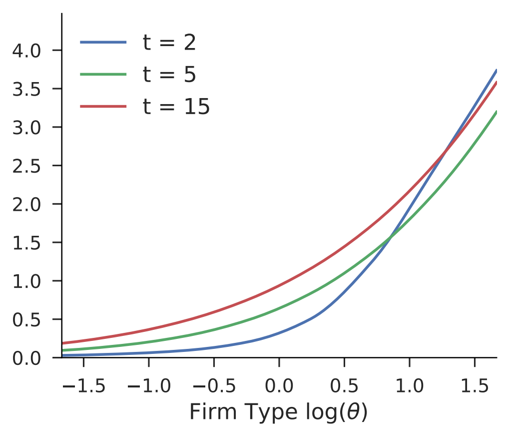
(a) Investments and Effort by Age



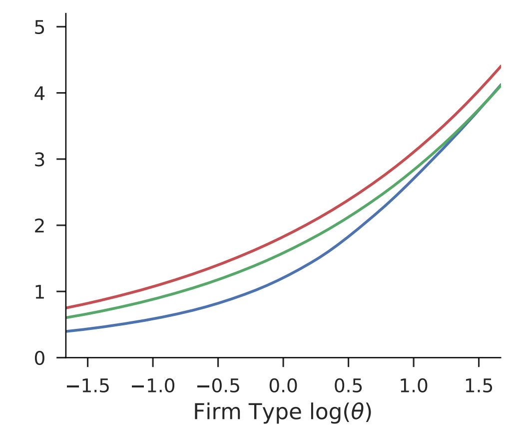
(b) Step Size and Profits by Age



(c) Effort by Type

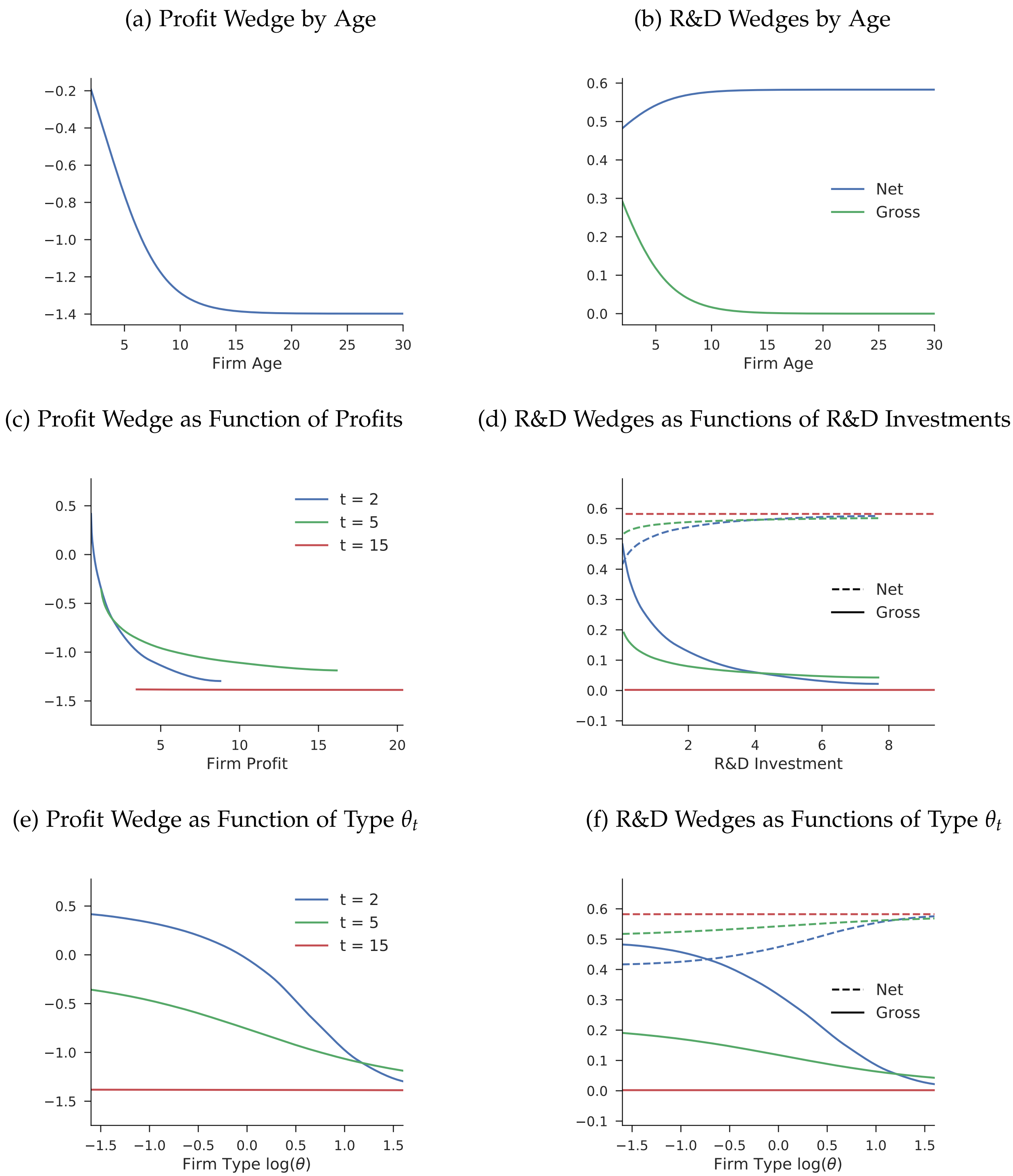


(d) R&D Investments by Type



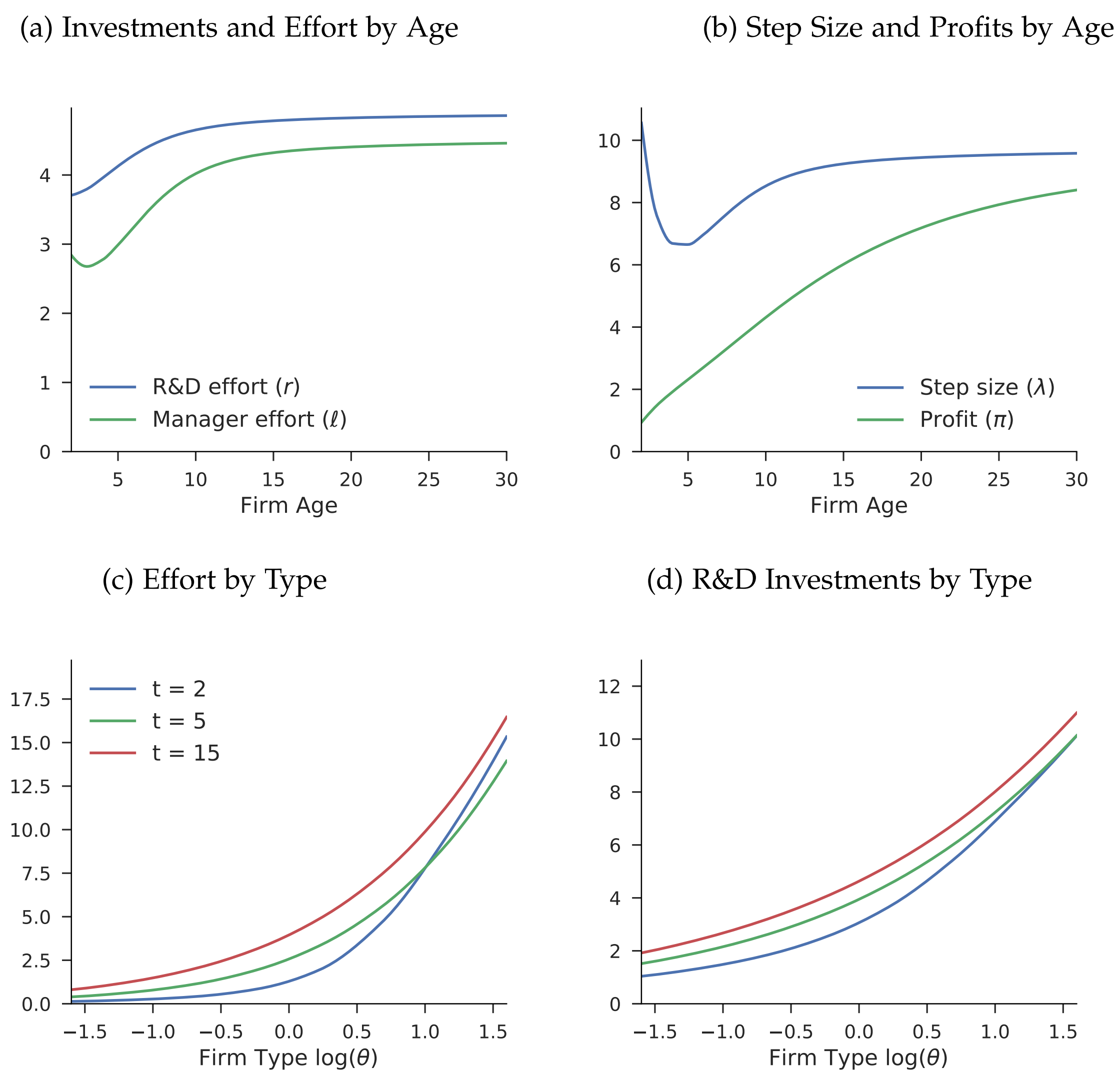
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S25: Optimal Profit and R&D Wedges, Overweighting Moment 1



Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

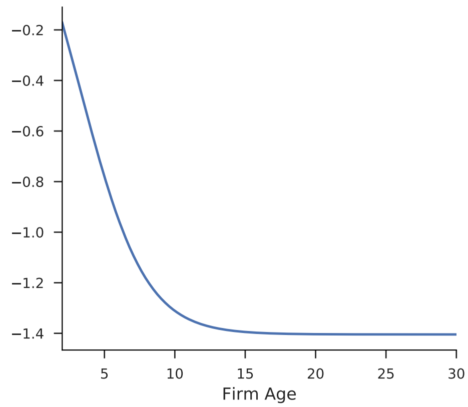
Figure S26: Optimal Allocations, Overweighting Moment 1



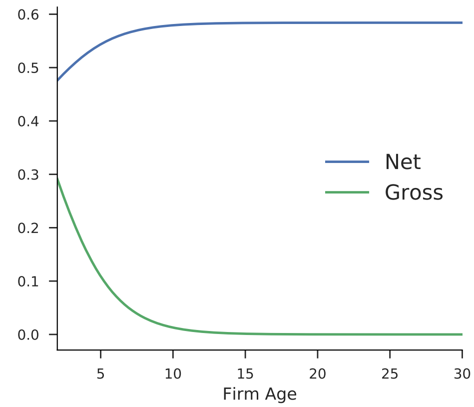
Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.

Figure S27: Optimal Profit and R&D Wedges, with $\eta = 1$

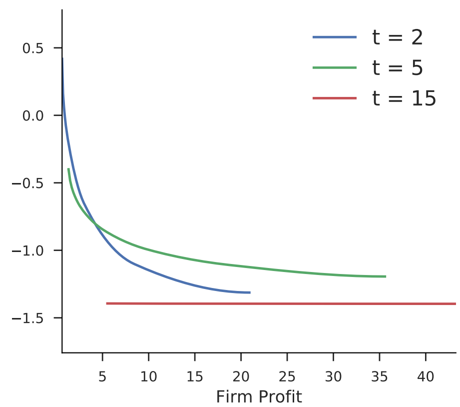
(a) Profit wedge by age



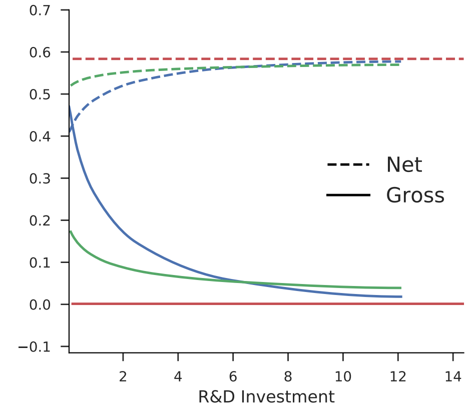
(b) (b) R&D wedges by age



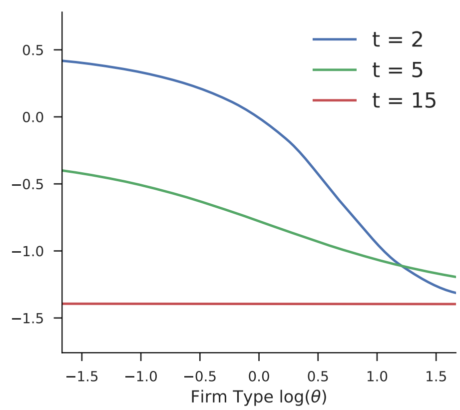
(c) Profit Wedge as Function of Profits



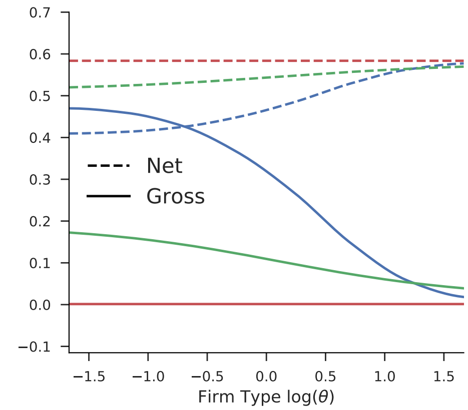
(d) R&D Wedges as Functions of R&D Investments



(e) Profit Wedge as Function of Type θ_t

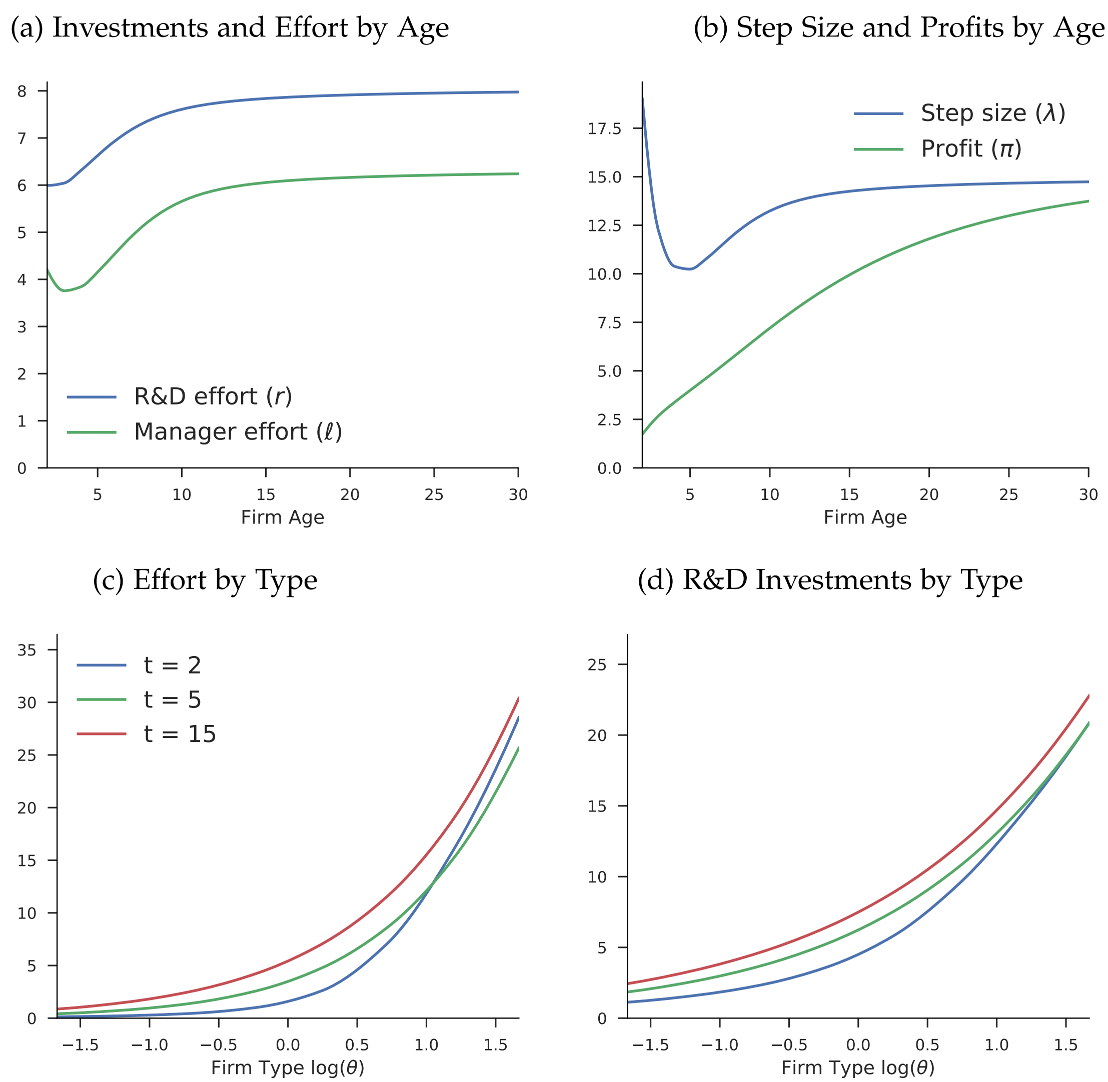


(f) R&D Wedges as Functions of Type θ_t



Notes: Panel (a) plots the average optimal profit wedge at different ages; Panel (b) plots the average optimal gross and net R&D wedges. Panels (c) and (d) plot, respectively, the optimal profit and R&D wedges for $t = 2, 5, 15$ for different levels of profits and R&D investments. Panels (e) and (f) plot the same wedges, but against firm productivity type θ_t.

Figure S28: Optimal Allocations for $\eta = 1$



Notes: The figure depicts the optimal allocations for different ages and types of firms. Panel (a) shows optimal investments in R&D and effort for different ages; panel (b) shows the resulting step size and profits by age. Panels (c) and (d) depict, respectively, the optimal R&D effort and R&D investments for firms of different types for ages 2, 5, and 15.