ONLINE APPENDICES FOR DISTORTED INNOVATION: DOES THE MARKET GET THE DIRECTION OF TECHNOLOGY RIGHT?

Appendix A: Additional Theoretical Results and Omitted Proofs

Derivation of Static Equilibrium

In this part of the Appendix, I provide a few expressions omitted from the text. First, the maximization of (9) gives the demand for machine varieties and resource inputs as

$$x_j(\nu) = \left[p_j \left(\frac{(1-\beta)\alpha}{(1+\mu_j)\psi} \right)^{\alpha} \left(\frac{1-\alpha}{q_j^R} \right)^{1-\alpha} \right]^{\frac{1}{\alpha\beta}} L_j, \tag{A1}$$

and

$$R_j = \left[p_j \left(\frac{(1-\beta)\alpha}{(1+\mu_j)\psi} \right)^{\alpha(1-\beta)} \left(\frac{1-\alpha}{q_j^R} \right)^{1-\alpha+\alpha\beta} \right]^{\frac{1}{\alpha\beta}} N_j L_j. \tag{A2}$$

Substituting these into (4), we obtain the levels of sectoral production as

$$Y_j = \left(\frac{(1-\beta)\alpha}{(1+\mu_j)\psi} \right)^{\frac{1-\beta}{\beta}} \left(\frac{1-\alpha}{q_j^R} \right)^{\frac{1-\alpha}{\alpha\beta}} p_j{}^{\frac{1-\alpha\beta}{\alpha\beta}} N_j L_j. \tag{A3}$$

Combining this expression for $j = 1, 2$ with (11) and rearranging yields (12) in the text.

I next show that the equilibrium characterized in the text is unique when $\delta\sigma < 1$ and there are always multiple (corner) equilibria when $\delta\sigma > 1$. Recall from footnote 9 that each scientist has a mass $\mathfrak{s} > 0$, and then we are taking the limit case where $\mathfrak{s} \to 0$. Then consider an allocation in which all researchers work in sector 2 (of course, the argument is analogous when they all work in sector 1). For this allocation to be an equilibrium, we need that switching to sector 1 is not profitable for an individual scientist. This requires

$$\tilde{\eta}_2 (\bar{S} - \mathfrak{s})^{\frac{\delta}{1-\delta}} \pi_2 \geq \tilde{\eta}_1 \mathfrak{s}^{\frac{\delta}{1-\delta}} \pi_1.$$

Rearranging this equation and using the equivalent conditions from the main text, it becomes

$$\frac{\tilde{\eta}_2}{\tilde{\eta}_1} \left(\frac{\bar{S} - \mathfrak{s}}{\mathfrak{s}} \right)^{\frac{\delta}{1-\delta}} \frac{\mu_2}{\mu_1} \left(\frac{1+\mu_2}{1+\mu_1} \right)^{-\frac{1}{\beta}} \left(\frac{q_2^R}{q_1^R} \right)^{-\frac{1-\alpha}{\alpha\beta}} p^{\frac{1}{\alpha\beta}} \left(\frac{L_2}{L_1} \right) \geq 1,$$

or substituting from (12) and recalling that $N_j = \tilde{\eta}_j S_j^{\frac{1}{1-\delta}}$, it is equivalent to

$$\frac{\eta_2}{\eta_1}\left(\frac{\gamma_2}{\gamma_1}\right)^{\frac{\varepsilon}{\sigma}}\frac{\mu_2}{\mu_1}\left(\frac{1+\mu_2}{1+\mu_1}\right)^{-\frac{\sigma-(1-\beta)}{\beta\sigma}}\left(\frac{q_2^R}{q_1^R}\right)^{-\frac{(\sigma-1)(1-\alpha)}{\alpha\beta\sigma}}\left(\frac{L_2}{L_1}\right)^{\frac{\sigma-1}{\sigma}}\left(\frac{\bar{S}-\mathfrak{s}}{\mathfrak{s}}\right)^{-\frac{1-\delta\sigma}{(1-\delta)\sigma}} \geq 1.$$

Now taking the limit $\mathfrak{s} \to 0$, we can see that this condition can never be satisfied when $\delta\sigma < 1$, since $\left(\frac{\bar{S}-\mathfrak{s}}{\mathfrak{s}}\right)^{-\frac{1-\delta\sigma}{(1-\delta)\sigma}} \to 0$, and hence the entire left-hand side limits to 0. Conversely, when $\delta\sigma > 1$, $\left(\frac{\bar{S}-\mathfrak{s}}{\mathfrak{s}}\right)^{-\frac{1-\delta\sigma}{(1-\delta)\sigma}} \to +\infty$ and thus the left-hand side limits to $+\infty$, ensuring that this condition is always satisfied as a strict inequality. This establishes that when $\delta\sigma < 1$, there are no corner equilibria and the interior equilibrium characterized in the text is unique. Conversely, when $\delta\sigma > 1$, corner allocations are always equilibria.

Derivation of Socially-Optimal Technology Ratio in the Static Model

The first-order conditions for the social planner in the static model can be written as:

$$\tilde{\eta}_1 S_1^{\frac{\delta}{1-\delta}}\left[\frac{d\ln Y}{dN_1}+\frac{d\ln E}{dN_1}\right] = \tilde{\eta}_2 S_2^{\frac{\delta}{1-\delta}}\left[\frac{d\ln Y}{dN_2}+\frac{d\ln E}{dN_2}\right].$$

From this expression, using the fact that $d\ln N_j = dN_j/N_j$ and substituting N_j for S_j from $N_j = \tilde{\eta}_j S_j^{\frac{1}{1-\delta}}$ (with $\eta_j \equiv \tilde{\eta}_j^{1-\delta}$), we get (19) in the main text.

Next, note that

$$\frac{d\ln Y}{d\ln Y_j} = \gamma_j\left(\frac{Y_j}{Y}\right)^{\frac{\varepsilon-1}{\varepsilon}} = \gamma_j^{\varepsilon} p_j^{1-\varepsilon},$$

where the second relationship exploits the representative household's utility maximization condition (recalling that the social planner does not directly manipulate prices). Moreover:

$$\begin{aligned}\frac{d\ln Y_2}{d\ln N_2} &= 1+\frac{\partial\ln Y_2}{\partial\ln p_2}\frac{d\ln p_2}{d\ln N_2}\\ &= 1+\frac{1-\alpha\beta}{\alpha\beta}\frac{d\ln p_2}{d\ln p}\frac{d\ln p}{d\ln N_2}\end{aligned}$$

$$\begin{aligned}\frac{d\ln Y_2}{d\ln N_1} &= \frac{\partial\ln Y_2}{\partial\ln p_2}\frac{d\ln p_2}{d\ln N_1}\\ &= \frac{1-\alpha\beta}{\alpha\beta}\frac{d\ln p_2}{d\ln p}\frac{d\ln p}{d\ln N_1}\end{aligned}$$

$$\begin{aligned}\frac{d\ln Y_1}{d\ln N_1} &= 1+\frac{\partial\ln Y_1}{\partial\ln p_1}\frac{d\ln p_1}{d\ln N_1}\\ &= 1+\frac{1-\alpha\beta}{\alpha\beta}\frac{d\ln p_1}{d\ln p}\frac{d\ln p}{d\ln N_1}\end{aligned}$$

$$\begin{aligned}\frac{d\ln Y_1}{d\ln N_2} &= \frac{\partial \ln Y_1}{\partial \ln p_1}\frac{d\ln p_1}{d\ln N_2} \\ &= \frac{1-\alpha\beta}{\alpha\beta}\frac{d\ln p_1}{d\ln p}\frac{d\ln p}{d\ln N_2}\end{aligned}$$

Finally, using the ideal price condition, (13),

$$\begin{aligned}\frac{dp_1/dp}{p_1} &= -\gamma_2^{\varepsilon}p^{-\varepsilon}p_1^{1-\varepsilon} \\ \frac{dp_2/dp}{p_2} &= \gamma_1^{\varepsilon}p^{\varepsilon-2}p_2^{1-\varepsilon}.\end{aligned}$$

Or

$$d\ln p_1/d\ln p = -\gamma_2^{\varepsilon}p_2^{1-\varepsilon} \text{ and } d\ln p_2/d\ln p = \gamma_1^{\varepsilon}p_1^{1-\varepsilon}.$$

Now combining these expressions, we have

$$\begin{aligned}\frac{d\ln Y}{d\ln N_1} &= \gamma_1^{\varepsilon}p_1^{1-\varepsilon}\left[1-\frac{(1-\alpha\beta)\gamma_2^{\varepsilon}p_2^{1-\varepsilon}}{\sigma}\right]+\gamma_2^{\varepsilon}p_2^{1-\varepsilon}\left[\frac{(1-\alpha\beta)\gamma_1^{\varepsilon}p_1^{1-\varepsilon}}{\sigma}\right] \\ &= \gamma_1^{\varepsilon}p_1^{1-\varepsilon},\end{aligned}$$

and

$$\begin{aligned}\frac{d\ln Y}{d\ln N_2} &= \gamma_1^{\varepsilon}p_1^{1-\varepsilon}\left[\frac{(1-\alpha\beta)\gamma_2^{\varepsilon}p_2^{1-\varepsilon}}{\sigma}\right]+\gamma_2^{\varepsilon}p_2^{1-\varepsilon}\left[1-\frac{(1-\alpha\beta)\gamma_1^{\varepsilon}p_1^{1-\varepsilon}}{\sigma}\right] \\ &= \gamma_2^{\varepsilon}p_2^{1-\varepsilon}.\end{aligned}$$

Combining these expressions, we obtain the desired result:

$$\frac{d\ln Y}{d\ln N_j} = \gamma_j^{\varepsilon}p_j^{1-\varepsilon} \text{ for } j=1,2.$$

Finally, using the same steps as in the previous subsection of the Appendix, we can show that when $\delta\sigma < 1$, the second-order conditions of the social planner's optimization problem are always satisfied in the interior allocation given by (16). Conversely, when $\delta\sigma > 1$, the interior allocation is not a local maximum, and instead there are two local maxima at the corners, with all scientists working in one of the two sectors. One of these two local maxima is the global maximum. Which one is preferred can be easily determined by using the expression for welfare derived in the next subsection of the Appendix and substituting for S_j in terms of N_j (once again from $N_j = \tilde{\eta}_j S_j^{\frac{1}{1-\delta}}$), and comparing the resulting expressions as $\mathfrak{s} \to 0$.

Measuring Externalities

In the theoretical analysis, I simplified the discussion by assuming that externalities are created directly by technology choices. This means that I need to convert existing externality estimates into those that appear in the form of the $\tilde{\tau}_j$ or τ_j variables. I now discuss how this can be done.

Automation: In the automation case, I follow Acemoglu, Manera and Restrepo's (2020) review of

the literature. The median estimate of quasi-rents (and thus pecuniary externalities) in labor income is about 15%. I combine this with Acemoglu and Restrepo's (2020) estimate of the effect of robot adoption on the employment in local labor markets (approximated by commuting zones in the US). Namely, let us equate automation technologies with N_2, and employment with L_1 (and L_2 can be capital or high-skilled labor working with automated technologies), and denote the working age population by Pop. Then we have

$$\begin{aligned}
\tilde{\tau}_2 &= -\frac{d\ln E}{d\ln N_2} = -\frac{d\ln E}{d\ln L_1}\frac{d\ln L_1}{d\ln N_2} \\
&= -\frac{d\ln E}{d\ln L_1}\frac{d\ln L_1}{dL_1}\frac{dL_1}{dL_1/\text{Pop}}\frac{dL_1/\text{Pop}}{dN_2}\frac{dN_2}{d\ln N_2} \\
&= -\frac{d\ln E}{d\ln L_1}\frac{\text{Pop}}{L_1}\frac{dL_1/\text{Pop}}{dN_2}N_2 \\
&= -(0.15)\times\frac{1}{0.63}\times(-0.39)\times 0.73 \\
&= 0.07,
\end{aligned}$$

where -0.15 is from Acemoglu, Manera and Restrepo's review of the literature, 0.63 is the employment to population ratio in the United States, averaged over the years 1990-2007 in Current Population Survey,[24] -0.39 is Acemoglu and Restrepo's (2020) estimate of the impact of one more robot per 1000 industrial workers on employment to population ratio, and 0.73 is their estimate of the stock of robots between 1993 and 2007. This number implies that a proportional increase in automation technology creates a 7% negative pecuniary externality on workers. I then convert this into τ_2 as described above.

Health care: In the health care case, the main distortion I focus on is differential markups, which does not need any conversion. Secondarily, I compute the externalities in terms of differences in quality-adjusted life year returns per one dollar of spending on technology between the preventative and high-tech/late-stage curative technologies (inclusive of R&D costs and usage costs). These numbers are therefore directly comparable to $\tilde{\tau}_j$'s in our model. Further details of medical procedures, drugs and technologies used in these computations and the studies from which the estimates are taken are provided in Appendix C.

Energy: In the energy case, I use estimates of the social cost of carbon. In this framework, carbon corresponds to the (suitably rescaled) resource input R_2 (identifying dirty technologies with sector 2). The social cost of carbon is in terms of the impact of one more metric ton of carbon emissions on consumption-equivalent welfare. The externality in the utility equation (1) is in terms of proportional effect on consumption. Therefore, I compute $\tilde{\tau}_2$ as follows:

$$\begin{aligned}
\tilde{\tau}_2 &= -\frac{d\ln E_2}{d\ln N_2} = -\frac{dE_2}{dR_2}\frac{R_2}{E_2}\frac{d\ln R_2}{d\ln N_2} \\
&= SCC\times\frac{\text{CO2 emission}}{\text{Energy Consumption}}\frac{d\ln R_2}{d\ln N_2},
\end{aligned}$$

where I am using the fact that the relevant consumption is total energy consumption and proxying $d\ln R_2/d\ln N_2 \simeq 1$.

[24] From FRED, https://fred.stlouisfed.org/series/EMRATIO

Welfare Comparisons

The welfare difference between the social optimum in the equilibrium can be written as

$$U^{SP} - U^{EQ} = \ln Y\left(n^{SP}\right) - \ln Y\left(n^{EQ}\right) + \ln E(n^{SP}) - \ln E(n^{EQ}),$$

where I am using the fact that all other endogenous variables are functions of n. The basic idea is to develop the approximation:

$$\Delta \ln Y^{EQ,SP} \equiv \ln Y\left(n^{SP}\right) - \ln Y\left(n^{EQ}\right) \simeq \frac{d \ln Y\left(n^{EQ}\right)}{dS}\left[S^{SP} - S^{EQ}\right], \tag{A4}$$

where S^{SP} is the allocation of scientists consistent with a technology ratio of n^{SP}, and S^{EQ} is the allocation of scientists implied by the technology ratio of n^{EQ}.

To do this, consider the impact of a change in the allocation of scientists from the equilibrium n^{EQ}, and let $S_1 = S$ and $S_2 = \bar{S} - S$. Then we can write:

$$\begin{aligned}
\frac{d \ln Y}{dS} &= \frac{d \ln Y}{d \ln N_1}\frac{d \ln N_1}{dS} - \frac{d \ln Y}{d \ln N_2}\frac{d \ln N_2}{dS} \\
&= \frac{d \ln Y}{d \ln N_1}\frac{dN_1}{dS}\frac{1}{N_1} - \frac{d \ln Y}{d \ln N_2}\frac{dN_2}{dS}\frac{1}{N_2} \\
&= \frac{1}{1-\delta}\left[\gamma_1^{\varepsilon} p_1^{1-\varepsilon}\tilde{\eta}_1 N_1^{-1} S_1^{\frac{\delta}{1-\delta}} - \gamma_2^{\varepsilon} p_2^{1-\varepsilon}\tilde{\eta}_2 N_2^{-1} S_2^{\frac{\delta}{1-\delta}}\right] \\
&= \frac{1}{1-\delta}\left[\gamma_1^{\varepsilon} p_1^{1-\varepsilon}\eta_1 N_1^{-(1-\delta)} - \gamma_2^{\varepsilon} p_2^{1-\varepsilon}\eta_2 N_2^{-(1-\delta)}\right] \\
&= \frac{1}{1-\delta}\left[\eta_1\gamma_1^{\varepsilon}\left[\gamma_1^{\varepsilon} + \gamma_2^{\varepsilon} p^{1-\varepsilon}\right]^{-1} N_1^{-(1-\delta)} - \eta_2\gamma_2^{\varepsilon}\left[\gamma_1^{\varepsilon} + \gamma_2^{\varepsilon} p^{1-\varepsilon}\right]^{-1} p^{1-\varepsilon} N_2^{-(1-\delta)}\right] \\
&= \frac{N_1^{-(1-\delta)}}{1-\delta}\left[\eta_1\gamma_1^{\varepsilon}\left[\gamma_1^{\varepsilon} + \gamma_2^{\varepsilon}\left(p\right)^{1-\varepsilon}\right]^{-1} - \eta_2\gamma_2^{\varepsilon}\left[\gamma_1^{\varepsilon} + \gamma_2^{\varepsilon}\left(p\right)^{1-\varepsilon}\right]^{-1} p^{1-\varepsilon} n^{-(1-\delta)}\right]
\end{aligned} \tag{A5}$$

Here, the third line simply uses the expressions for dN_j/dS_j from the static innovation possibilities frontier (5), while the fourth line uses the same transformation as in the text: $N_j = \tilde{\eta}_j S_j^{\frac{1}{1-\delta}}$ and $\eta_j \equiv \tilde{\eta}_j^{1-\delta}$. The penultimate line uses the ideal price condition (13) to substitute p_1 and p_2 in terms of the relative price p. The final line simply factors out N_1.

Hence,

$$\begin{aligned}
\Delta \ln Y^{EQ,SP} &\simeq \frac{d\ln Y}{dS}\left(S^{SP}-S^{EQ}\right)\\
&= \frac{d\ln Y}{dS}\left(\frac{\bar{S}}{1+\frac{\eta_1}{\eta_2}\left(n^{SP}\right)^{1-\delta}}-\frac{\bar{S}}{1+\frac{\eta_1}{\eta_2}\left(n^{EQ}\right)^{1-\delta}}\right)\\
&= \frac{\left(N_1^{EQ}\right)^{-(1-\delta)}}{1-\delta}\begin{pmatrix}\eta_1\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\\ -\eta_2\gamma_2^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\left(p^{EQ}\right)^{1-\varepsilon}\left(n^{EQ}\right)^{-(1-\delta)}\end{pmatrix}\left(S^{SP}-S^{EQ}\right)\\
&= \frac{\left(N_1^{EQ}\right)^{-(1-\delta)}}{1-\delta}\eta_1\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\left(1-\frac{\eta_2}{\eta_1}\left(\frac{\gamma_2}{\gamma_1}\right)^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\left(n^{EQ}\right)^{-(1-\delta)}\right)\left(S^{SP}-S^{EQ}\right)\\
&= \frac{\left(N_1^{EQ}\right)^{-(1-\delta)}}{1-\delta}\eta_1\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]\left(S^{SP}-S^{EQ}\right)\\
&= \frac{\left(N_1^{EQ}\right)^{-(1-\delta)}}{1-\delta}\eta_1\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]\left(\frac{\bar{S}}{1+\frac{\eta_1}{\eta_2}\left(n^{SP}\right)^{1-\delta}}-\frac{\bar{S}}{1+\frac{\eta_1}{\eta_2}\left(n^{EQ}\right)^{1-\delta}}\right)\\
&= \frac{1}{1-\delta}\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]\left(\frac{1+\frac{\eta_1}{\eta_2}\left(n^{EQ}\right)^{1-\delta}}{1+\frac{\eta_1}{\eta_2}\left(n^{SP}\right)^{1-\delta}}-1\right).
\end{aligned}$$

In these derivations, I have used the following steps. The second line is from (A4), while the third line substitutes from (A5). The fourth line factors out $\eta_1\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}$. The fifth line uses the expressions for p^{EQ} and n^{EQ} from (12) and (16). The sixth line uses the fact that from (5), the equilibrium and socially-optimal allocations of scientists have to satisfy

$$S^{EQ}=\frac{\bar{S}}{1+\frac{\eta_1}{\eta_2}\left(n^{EQ}\right)^{1-\delta}} \text{ and } S^{SP}=\frac{\bar{S}}{1+\frac{\eta_1}{\eta_2}\left(n^{SP}\right)^{1-\delta}}.$$

The seventh line then substitutes for

$$N_1^{EQ}=\left(\frac{\bar{S}}{\frac{1}{\eta_1}+\frac{\left(n^{EQ}\right)^{1-\delta}}{\eta_2}}\right)^{\frac{1}{1-\delta}}, \tag{A6}$$

and cancels out terms.

Here everything is a function of n^{EQ} and parameters.

With no markup differences, it can be verified that $\Delta \ln Y^{EQ,SP}=0$ (which also follows from an application of the envelope theorem). Note, in particular, that the terms in square brackets are equal to

zero. Therefore, with no markup differences (as in the automation and energy cases), we have

$$\begin{aligned} U^{SP} - U^{EQ} &\simeq \ln E(n^{SP}) - \ln E(n^{EQ}) \\ &= (\tilde{\tau}_1 + \tilde{\tau}_2) \ln \left(\frac{N_1^{EQ}}{N_1^{SP}} \right) + \tilde{\tau}_2 \ln \left(\frac{n^{EQ}}{n^{SP}} \right). \end{aligned}$$

This can be computed given the expression for N_1^{EQ} in (A6), and its analogue for the socially-optimal level of technology:

$$N_1^{SP} = \left(\frac{\bar{S}}{\frac{1}{\eta_1} + \frac{(n^{SP})^{1-\delta}}{\eta_2}} \right)^{\frac{1}{1-\delta}}.$$

Hence,

$$\frac{N_1^{EQ}}{N_1^{SP}} = \left(\frac{1 + \frac{\eta_1}{\eta_2} \left(n^{SP}\right)^{1-\delta}}{1 + \frac{\eta_1}{\eta_2} \left(n^{EQ}\right)^{1-\delta}} \right)^{\frac{1}{1-\delta}}.$$

To proxy for the ratio N_1^{EQ}/N_1^{SP} we only need an estimate of η_2/η_1. This ratio can be obtained by using the expressions for (12), (16), and (A3). Combining these equations, we obtain:

$$\frac{\eta_2}{\eta_1} = \left(n^{EQ}\right)^{1-\delta} \left(\frac{\mu_2}{\mu_1} \right)^{-1} \left(\frac{1+\mu_2}{1+\mu_1} \right) \left(\frac{p_2 Y_2}{p_1 Y_1} \right)^{-1},$$

where $p_2 Y_2 / p_1 Y_1$ is the relative output of the two sectors.

When there are markup differences and no externalities (as in our baseline health care application), then

$$U^{SP} - U^{EQ} \simeq \Delta \ln Y^{EQ,SP}$$

as derived above. This can be computed if we can also compute p^{EQ} and have an estimate for γ_2/γ_1. In the health care application, we have $\alpha = 1$ and there is no specialized labor, so the same health care labor forces allocated between the two technologies, which implies

$$\frac{w_2^{EQ}}{w_1^{EQ}} = \left(p^{EQ}\right)^{\frac{1}{\beta}} \left(\frac{1+\mu_2}{1+\mu_1} \right)^{-\frac{1-\beta}{\beta}} \left(n^{EQ}\right)^{\frac{\sigma-1}{\sigma}} = 1.$$

Given markups and n^{EQ}, this equation gives p^{EQ}. To obtain an estimate for γ_2/γ_1, note first that $\gamma_1 + \gamma_2 = 1$, and thus $\gamma_2/\gamma_1 = (1-\gamma_1)/\gamma_1$. Moreover,

$$p = \frac{1-\gamma_1}{\gamma_1} \left(\frac{Y_2}{Y_1} \right)^{-\frac{1}{\varepsilon}},$$

which can be rearranged to yield

$$\gamma_1 = \left[1 + \left(p^{EQ}\right)^{\frac{\varepsilon-1}{\varepsilon}} \left(\frac{p_2 Y_2}{p_1 Y_1} \right)^{\frac{1}{\varepsilon}} \right]^{-1},$$

which again uses an estimate of the relative output of the two sectors.[25]

This discussion clarifies that to compute the welfare losses from innovation distortions we need two more numbers in each applications: n^{EQ} and $(p_2Y_2)/(p_1Y_1)$. We use the following estimates for these quantities:

- *Automation*: n^{EQ} is taken as the ratio of the total number of automation patents to total non-automation patents across all countries in 2005 from Acemoglu and Restrepo (2022). This gives $n^{EQ} = 0.15$. I also set $(p_2Y_2)/(p_1Y_1) = 0.38$ on the basis of the model-based inference in Acemoglu, Manera and Restrepo (2020), which yields that about 28% of tasks are automated in the US economy.

- *Health care*: n^{EQ} is taken as the ratio of the sum of the discounted stock of curative patents to that of the sum of preventative patents across countries, which gives $n^{EQ} = 16.2$. I proxy p_1Y_1 as total spending on ambulatory health services and social assistance in 2020 from the U.S. Census Bureau Service Annual Survey (SAS) and total spending on diagnostic substances and biological product manufacturing in 2020 from the U.S. Census Bureau and Annual Survey of Manufactures (ASM). I set p_2Y_2 equal to the 2020 revenues of the same industries classified as curative in Appendix C. These revenues are also taken from the ASM. The resulting ratio is $(p_2Y_2)/(p_1Y_1) = 0.13$.[26]

- Energy: $n^{EQ} = 2.20$ is taken as the sum of the stock of dirty patents to that of the stock of clean patents across all countries in 2005, from Aghion et al. (2016). In addition, $(p_2Y_2)/(p_1Y_1)$ is proxied by the ratio of the revenue of renewable energy to that of non-renewable energy, where revenue is calculated as the product of average wholesale electricity price of an energy source and its primary energy consumption from the EIA Monthly Energy Review. The resulting ratio is $(p_2Y_2)/(p_1Y_1) = 3.08$.[27]

[25] An alternative approximation for welfare in this case can be derived by taking a first-order Taylor approximation in terms of deviations between technology ratios and then substituting out some of the technology terms by using the same equilibrium relationship we used in the welfare computations, in particular, $\frac{\eta_1}{\eta_2}\left(n^{EQ}\right)^{1-\delta} = \left(\frac{\gamma_2}{\gamma_1}\right)^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]^{-1}$. This gives the following approximation for output differences between the optimal and equilibrium allocations in the presence of markup differences: $\gamma_1^{\varepsilon}\left(\gamma_1^{\varepsilon}+\gamma_2^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\right)^{-1}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]\left(\frac{\left(\frac{\gamma_2}{\gamma_1}\right)^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]^{-1}}{1+\left(\frac{\gamma_2}{\gamma_1}\right)^{\varepsilon}\left(p^{EQ}\right)^{1-\varepsilon}\left[1-\frac{\mu_1}{\mu_2}\frac{1+\mu_2}{1+\mu_1}\right]^{-1}}\right)$. This approximation removes the need to separately estimate η_2/η_1. In practice, the two expressions give very similar estimates of welfare costs of distorted technology in the health care case.

[26] More specifically, the preventative categories are: NAICS 621 (Ambulatory health), NAICS 624 (Social assistance), NAICS 325413 (In-vitro diagnostic substances manufacturing), and NAICS 325414 (Biological product manufacturing). The curative categories are: NAICS 325412 (Pharmaceutical preparation manufacturing), NAICS 334510 (Electromedical and electrotherapeutic apparatus manufacturing), NAICS 339112 (Surgical and medical instrument manufacturing), and NAICS 339113 (Surgical appliance and supplies manufacturing). See Appendix C for details. The SAS and ASM data can be accessed at www.census.gov/programs-surveys/sas/data/tables.html and www.census.gov/programs-surveys/asm/data/tables.html, respectively.

[27] Wholesale prices are from the United States Energy Information Agency (EIA) Power Operations Report (see www.eia.gov/energyexplained/us-energy-facts/) and energy consumption data are from the EIA Monthly Energy Review (www.eia.gov/todayinenergy/detail.php?id=45436).

Dynamic Model

In this part of the Appendix, I provide a few more details about the dynamic framework provided in the text. First recall that when a scientist invents a new machine for sector $j \in \{1,2\}$, she receives the net present discounted value of future profits from the sale of this machine, given by

$$V_j(t) = \int_t^{\infty} e^{-\int_t^{t'} r(t'')dt''} \pi_j(t')dt', \tag{A7}$$

where $r(t)$ is the market interest rate at time t, and $\pi_j(t)$ is the common profit that all machines for sector $j \in \{1,2\}$ will make at time t.

The representative household's optimization problem implies that the growth rate of consumption has to satisfy

$$\frac{\dot{C}(t)}{C(t)} = r(t) - \varrho, \tag{A8}$$

as well as a standard transversality condition, which requires the net present discounted value of current and future machine varieties to be finite (see Acemoglu, 2002).

In BGP, consumption has to grow at a constant rate, and thus the interest rate will be constant. Therefore, we have

$$V_j = \frac{\pi_j}{r} \text{ for } j = 1,2.$$

Using these expressions for the two sectors and combining them with the equilibrium allocation of scientists, we obtain (16), as claimed in the text, which also establishes Proposition 2. The proof for Proposition 3 follows the analysis in Acemoglu (2002) closely and I do not present it here to avoid repetition.

I next consider the socially optimal choice of technology in this dynamic framework. Once again, assuming that the social planner only controls the allocation of scientists, this problem can be written as

$$\max_{[S(t),N_1(t),N_2(t)]_0^{\infty}} \int_0^{\infty} e^{-\varrho t} U[N_1(t), N_2(t)]dt$$

subject to

$$\dot{N}_1(t) = \eta_1 N_1(t)^{(1+\delta)/2} N_2(t)^{(1-\delta)/2} S(t) \tag{A9}$$

and

$$\dot{N}_2(t) = \eta_2 N_1(t)^{(1-\delta)/2} N_2(t)^{(1+\delta)/2} [\bar{S} - S(t)]. \tag{A10}$$

Here $U[N_1(t), N_2(t)] = \ln C(t) + \ln E(t)$ is the level of utility at time t, inclusive of externalities, given the vector of technologies (state variables), $N_1(t)$ and $N_2(t)$. This expression exploits the fact that the level of final good production and hence consumption only depend on the current state of technologies. (All other endogenous variables, and in particular prices of the intermediates, $p_1(t)$ and $p_2(t)$, are solved out as in the equilibrium allocation in the text).

Suppressing time dependence when this will cause no confusion and assigning co-state variables λ_1 and λ_2 to (A9) and (A10), the necessary condition from the maximum principle applied to this optimal

control problem yields:

$$\lambda_1\eta_1 N_1^{\frac{1+\delta}{2}} N_2^{\frac{1-\delta}{2}} - \lambda_2\eta_2 N_1^{\frac{1-\delta}{2}} N_2^{\frac{1+\delta}{2}} \begin{cases} > 0 & \Longrightarrow \quad S = \bar{S} \\ = 0 & \Longrightarrow \quad S \in [0, \bar{S}] \\ < 0 & \Longrightarrow \quad S = 0 \end{cases} . \tag{A11}$$

Therefore, just like in the equilibrium, the social planner's solution leads to a bang-bang solution. Moreover, for an interior BGP, we need scientists to be assigned to both sectors, and thus this expression should be equal to zero. Hence, in an interior BGP, we must have:

$$\lambda_1\eta_1 = \lambda_2\eta_2 n^{\delta}. \tag{A12}$$

In order to characterize the socially-optimal technology choices, we need to know the values and evolution of the co-state variables, which are given by the following two differential equations:

$$\begin{aligned}
\varrho\lambda_1 - \dot{\lambda}_1 &= \frac{dU}{dN_1} + \frac{1+\delta}{2}\lambda_1\eta_1 \left(\frac{N_2}{N_1}\right)^{\frac{1-\delta}{2}} S + \frac{1-\delta}{2}\lambda_2\eta_2 \left(\frac{N_2}{N_1}\right)^{\frac{1+\delta}{2}} \left[\bar{S} - S\right] \\
&= \frac{dU}{dN_1} + \frac{1+\delta}{2}\lambda_1\eta_1 n^{\frac{1-\delta}{2}} S + \frac{1-\delta}{2}\lambda_2\eta_2 n^{\frac{1+\delta}{2}} \left(\bar{S} - S\right) \\
&= \frac{dU}{dN_1} + \lambda_1\eta_1 n^{\frac{1-\delta}{2}} \left[\frac{1+\delta}{2} S + \frac{1-\delta}{2}\left(\bar{S} - S\right)\right], \\
\varrho\lambda_2 - \dot{\lambda}_2 &= \frac{dU}{dN_2} + \frac{1-\delta}{2}\lambda_1\eta_1 \left(\frac{N_2}{N_1}\right)^{-\frac{1+\delta}{2}} S + \frac{1+\delta}{2}\lambda_2\eta_2 \left(\frac{N_2}{N_1}\right)^{-\frac{1-\delta}{2}} \left[\bar{S} - S\right] \\
&= \frac{dU}{dN_2} + \frac{1-\delta}{2}\lambda_1\eta_1 n^{-\frac{1+\delta}{2}} S + \frac{1+\delta}{2}\lambda_2\eta_2 n^{-\frac{1-\delta}{2}} \left(\bar{S} - S\right) \\
&= \frac{dU}{dN_2} + \lambda_2\eta_2 n^{-\frac{1-\delta}{2}} \left[\frac{1-\delta}{2} S + \frac{1+\delta}{2}\left(\bar{S} - S\right)\right].
\end{aligned}$$

In BGP, we need $\dot{\lambda}_1 = \dot{\lambda}_2 = 0$, and hence

$$\begin{aligned}
\lambda_1 &= \frac{1}{\varrho}\left(\frac{dU}{dN_1} + \lambda_1\eta_1 n^{\frac{1-\delta}{2}} \left[\frac{1+\delta}{2} S + \frac{1-\delta}{2}\left(\bar{S} - S\right)\right]\right) \\
&= \frac{\frac{dU}{dN_1}}{\varrho - \eta_1 n^{\frac{1-\delta}{2}} \left[\frac{1+\delta}{2} S + \frac{1-\delta}{2}\left(\bar{S} - S\right)\right]}, \text{ and} \\
\lambda_2 &= \frac{1}{\varrho}\left(\frac{dU}{dN_2} + \lambda_2\eta_2 n^{-\frac{1-\delta}{2}} \left[\frac{1-\delta}{2} S + \frac{1+\delta}{2}\left(\bar{S} - S\right)\right]\right) \\
&= \frac{\frac{dU}{dN_2}}{\varrho - \eta_2 n^{-\frac{1-\delta}{2}} \left[\frac{1-\delta}{2} S + \frac{1+\delta}{2}\left(\bar{S} - S\right)\right]}.
\end{aligned}$$

Moreover, the scientist allocation has to satisfy the BGP condition:

$$\frac{S}{\bar{S} - S} = \frac{\eta_2}{\eta_1} n^{-(1-\delta)}. \tag{A13}$$

Substituting the values of the co-state variables into (A12), we obtain

$$\eta_1 \frac{\frac{d\ln Y}{d\ln N_1} + \frac{d\ln E}{d\ln N_1}}{\varrho - \eta_1 n^{\frac{1-\delta}{2}} \left[\frac{1+\delta}{2} S + \frac{1-\delta}{2}\left(\bar{S} - S\right)\right]} = \eta_2 n^{-(1-\delta)} \frac{\frac{d\ln Y}{d\ln N_2} + \frac{d\ln E}{d\ln N_2}}{\varrho - \eta_2 n^{-\frac{1-\delta}{2}} \left[\frac{1-\delta}{2} S + \frac{1+\delta}{2}\left(\bar{S} - S\right)\right]}. \tag{A14}$$

This condition is different from (19) because the social planner takes into account the knowledge spillovers the two sectors create, which have differential effects depending on the relative technology ratio. In the special case where $\delta = 1$, we can combine (A13) and (A14) to show that these differential knowledge spillovers cancel out and we end up with the following condition for the socially-optimal BGP technology ratio:

$$\eta_1 \left[\frac{d\ln Y}{d\ln N_1} + \frac{d\ln E}{d\ln N_1}\right] = \eta_2 \left[\frac{d\ln Y}{d\ln N_2} + \frac{d\ln E}{d\ln N_2}\right],$$

which is identical to (20) when $\delta = 1$, and thus the same n^{SP} in (20) in the text characterizes the socially-optimal BGP technology ratio. The general case where $\delta < 1$ captures the same economic forces I emphasized in the text, but does not admit a closed-form solution for the technology ratio.

Appendix B: Robustness Checks

This part of the Appendix provides robustness checks on the regression results reported in Table 2 in the text. Table B1 considers variations for the automation regressions, Table B2 presents robustness checks for the the regressions on the relationship between medical research and disease burden, and finally Table B3 focuses on the relationship between fuel prices and direction of innovation in automobiles. The results of all three tables are discussed in the text.

The formulae for the path dependence parameter δ and the elasticity of substitution σ in the various tables and columns are:

Table B1, columns 1-8 (Long-run effects from relative market sizes)

$$\hat{\delta} = \max\{0, 1 - \hat{\rho}\}, \text{ and}$$

$$\hat{\sigma} = \frac{1 + \hat{\chi} - \hat{\delta}}{1 + \hat{\delta}\hat{\chi} - \hat{\delta}}.$$

Table B1, columns 9-10 (Long-run effects from relative prices and with spillovers)

$$\hat{\delta} = \max\{0, 1 - \hat{\rho} - \hat{\rho}_{\text{spillover}}\}, \text{ and}$$

$$\hat{\sigma} = \frac{2\hat{\chi} + \hat{\delta} - 1 - \hat{\chi}\hat{\delta}}{\hat{\chi} + \hat{\delta} - 1},$$

where $\hat{\rho}_{\text{spillover}}$ is the coefficient on the (relative technology) spillover term.

Table B2, columns 1-12 (Long-run effects from relative market sizes)

$$\hat{\delta} = \max\{0, 1 - \hat{\rho}\}, \text{ and}$$

$$\hat{\sigma} = \frac{1 + \hat{\chi} - \hat{\delta}}{1 + \hat{\delta}\hat{\chi} - \hat{\delta}}.$$

Table B3, columns 1-10 (Long-run effects from relative input prices)

$$\hat{\delta} = \max\{0, 1 - \hat{\rho}\}, \text{ and}$$

$$\hat{\sigma} = \frac{\alpha\beta\hat{\chi} - (1 - \hat{\delta})(1 - \alpha)}{\alpha\beta\hat{\delta}\hat{\chi} - (1 - \hat{\delta})(1 - \alpha)}.$$

Table B3, columns 11-12 (Long-run effects from relative input prices and with spillovers)

$$\hat{\delta} = \max\{0, 1 - \hat{\rho} - \hat{\rho}_{\text{spillover}}\}, \text{ and}$$

$$\hat{\sigma} = \frac{\alpha\beta\hat{\chi} - (1 - \hat{\delta})(1 - \alpha)}{\alpha\beta\hat{\delta}\hat{\chi} - (1 - \hat{\delta})(1 - \alpha)}.$$

Table B1: Robustness for Automation Application

LHS	$\ln(x)$	$\ln(x)$	$\ln(1+x)$	$\ln(1+x)$	$\text{asinh}(x)$	$\text{asinh}(x)$	$\ln(x)$	$\ln(x)$	$\ln(1+x)$	$\ln(1+x)$
Samples	Full	Full	Full	Full	Full	Full	OECD	OECD	Firm-Level	Firm-Level
Frequency	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year	5-year	5-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A Parameters Estimated from Regressions										
Initial Relative Stock: $\hat{\rho}$	0.78	0.81	0.41	0.67	0.26	0.52	0.19	0.46	0.83	0.83
	(0.13)	(0.12)	(0.12)	(0.14)	(0.09)	(0.12)	(0.06)	(0.17)	(0.03)	(0.03)
Initial Shifter: $\hat{\chi}$	0.84	1.11	1.00	1.07	0.63	1.04	0.38	1.07	1.66	2.06
	(0.39)	(0.38)	(0.51)	(0.53)	(0.29)	(0.41)	(0.20)	(0.35)	(0.69)	(0.85)
Changes in Shifter: $\hat{\lambda}$	1.11	2.12	1.34	2.55	0.42	1.72	0.14	1.06	-1.58	-0.52
	(0.54)	(0.67)	(0.51)	(0.90)	(0.29)	(0.63)	(0.16)	(0.65)	(0.77)	(0.95)
Spillovers:									-0.30	-0.23
									(0.15)	(0.21)
Observations	232	125	345	165	345	165	149	78	3,459	3,447
Country covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Firm fixed effects									Yes	Yes
Industry × Year fixed effects									Yes	Yes
Country × Year fixed effects										Yes
Panel B Implied Parameters										
Long-run Effects	1.09	1.36	2.43	1.61	2.40	2.02	2.00	2.34	3.13	3.43
$\hat{\delta}$	0.22	0.19	0.59	0.33	0.74	0.48	0.81	0.54	0.47	0.40
$\hat{\sigma}$	1.68	1.88	1.41	1.70	1.22	1.53	1.15	1.47	1.78	1.85
$\hat{\varepsilon}$	4.06	4.96	2.86	4.13	2.01	3.37	1.66	3.12	4.50	4.81
$\hat{\delta}\hat{\sigma}$	0.38	0.35	0.83	0.57	0.91	0.74	0.93	0.80	0.84	0.74
Panel C Equilibrium and Welfare Comparison (Baseline: $\tilde{\tau} = 0.07$)										
n^{SP}/n^{EQ}	0.83	0.82	0.56	0.76	0.40	0.66	0.34	0.60	0.47	0.61
$U^{SP} - U^{EQ}$	0.01	0.01	0.03	0.01	0.04	0.02	0.05	0.02	0.03	0.02
Panel D Equilibrium and Welfare Comparison (Alternative: $\tilde{\tau} = 0.03$)										
n^{SP}/n^{EQ}	0.91	0.90	0.75	0.87	0.64	0.82	0.58	0.78	0.69	0.78
$U^{SP} - U^{EQ}$	0.002	0.002	0.01	0.003	0.01	0.004	0.01	0.01	0.01	0.01

Notes: This table presents regression estimates (Panel A), implied parameter values (Panel B) and implied distortions and welfare results (Panels C and D) for the automation application. Regressions are estimated with ordinary least squares and heteroscedasticity-robust standard errors clustered at country-level are presented in parentheses. All regressions are weighted by manufacturing employment in 1990. The dependent variable is relative number of newly granted patents for automation technologies relative to other utility patents divided by relative stock of patents related to automation relative to other utility patents (in logs, unless otherwise indicated). Shifters are expected 20-year level and change of the ratio of workers above the age of 56 to workers between 21 and 55 (in logs). Country covariates, included in columns 1-4, are region dummies, and the 1990 values of log GDP per capita, log of population, average years of schooling and the ratio of workers above 56 to workers aged 21 in 1990 interacted with period dummies. Columns 1 and 2 replicate the specifications from Table 2. Columns 3 and 4 use $\ln(1+x)$, while columns 5 and 6 use the inverse hyperbolic sine transformation. Columns 7 and 8 are for the OECD sample (with $\ln x$ as in our main specifications). Columns 9 and 10 report estimates from Dechezleprêtre et al.'s (2022) firm-level data, using a sample of firms with at least four automation patents. These regressions also include spillovers from country-level relative stock of knowledge. Column 9 controls for firm fixed effects and industry by year fixed effects, while column 10 additionally includes country by time fixed effects. The parameters δ and σ in these two columns are computed using the equations with spillovers provided above. Panel C uses 15% quasi-rents for workers, and Panel D uses 7.5% quasi-rents.

Table B2: Robustness for Health Application

LHS	$\ln(x)$	$\ln(x)$	$\ln(1+x)$	$\ln(1+x)$	$\text{asinh}(x)$	$\text{asinh}(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$
Samples	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full	US	US
Frequency	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A Parameters Estimated from Regressions												
Initial Relative Stock: $\hat{\rho}$	0.93	1.11	0.36	0.83	0.38	0.84	0.45	0.48	0.93	1.09	0.94	1.27
	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.10)	(0.12)
Initial Shifter: $\hat{\chi}$	0.10	0.14	0.05	0.11	0.07	0.13	0.07	0.10	0.11	0.14	0.32	0.26
	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.12)	(0.10)
Changes in Shifter: $\hat{\lambda}$	-0.004	0.001	-0.01	0.002	-0.01	0.002	-0.04	-0.06	-0.004	0.01	0.26	0.12
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.17)	(0.07)
Observations	55,699	37,389	75,399	44,569	75,399	44,569	55,702	37,394	55,625	37,358	1,243	741
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Disease fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			Yes	Yes
Period-Country fixed effects									Yes	Yes		
Period-Disease fixed effects									Yes	Yes		
Panel B Implied Parameters												
Long-run Effects	0.11	0.14	0.15	0.13	0.18	0.16	0.16	0.22	0.11	0.14	0.34	0.26
$\hat{\delta}$	0.07	0.00	0.64	0.17	0.62	0.16	0.55	0.52	0.07	0.00	0.06	0.00
$\hat{\sigma}$	1.10	1.14	1.05	1.11	1.06	1.13	1.07	1.09	1.11	1.14	1.31	1.26
$\hat{\varepsilon}$	1.18	1.26	1.09	1.19	1.11	1.24	1.12	1.17	1.19	1.26	1.57	1.47
$\hat{\delta}\hat{\sigma}$	0.08	0.00	0.67	0.19	0.66	0.18	0.58	0.57	0.08	0.00	0.08	0.00
Panel C Equilibrium and Welfare Comparison (Markups Only)												
n^{SP}/n^{EQ}	0.43	0.45	0.11	0.39	0.11	0.38	0.17	0.17	0.43	0.45	0.37	0.42
$U^{SP} - U^{EQ}$	0.06	0.05	0.15	0.06	0.15	0.07	0.12	0.12	0.06	0.05	0.06	0.06
Panel D Equilibrium and Welfare Comparison (Externalities Only)												
n^{SP}/n^{EQ}	0.58	0.59	0.24	0.54	0.24	0.53	0.31	0.32	0.58	0.59	0.52	0.56
$U^{SP} - U^{EQ}$	0.18	0.17	0.48	0.21	0.47	0.21	0.39	0.38	0.18	0.17	0.22	0.19

Notes: This table presents regression estimates (Panel A), implied parameter values (Panel B) and implied distortions and welfare results (Panels C and D) for the health application. Regressions are unweighted and estimated with ordinary least squares and heteroscedasticity-robust standard errors clustered at country-level are presented in parentheses. Observations are at the country-disease-period level. The dependent variable is relative number of new medical articles for each disease divided by relative stock of medical articles for that disease (in logs, unless otherwise indicated). Columns 1 and 2 replicate the main specifications from Table 2. Columns 3 and 4 use $\ln(1+x)$, while columns 5 and 6 use the inverse hyperbolic sine transformation. Columns 7 and 8 drop the country fixed effects, while columns 9 and 10 include period times country and period times disease fixed effects. Columns 11 and 12 focus on just the US observations. Panel C considers the implications of markup differences, and Panel D depicts the implications of an externality estimate based on the shortfall of quality-adjusted life year gains from curative vs. preventative technologies.

Table B3: Robustness for Energy Application

LHS	$\ln(1+x)$	$\ln(1+x)$	asinh(x)	asinh(x)	$\ln(1+x)$	$\ln(1+x)$	$\ln(1+x)$	$\ln(1+x)$	$\ln(1+x)$	$\ln(1+x)$
Frequency	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A Parameters Estimated from Regressions										
Initial Relative Stock: $\hat{\rho}$	0.81	0.86	0.82	0.86	0.58	0.51	0.81	0.86	0.81	0.86
	(0.03)	(0.04)	(0.02)	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.03)	(0.04)
Initial Shifter: $\hat{\chi}$	-1.52	-1.06	-1.99	-1.44	-0.07	-0.48	-1.51	-2.02	-1.50	-1.14
	(0.29)	(0.66)	(0.36)	(0.81)	(0.09)	(0.14)	(0.28)	(0.40)	(0.29)	(0.67)
Changes in Shifter: $\hat{\lambda}$	-0.45	1.12	-0.61	1.38	0.21	1.16	-0.47	-0.28	-0.43	1.09
	(0.20)	(0.82)	(0.26)	(1.00)	(0.13)	(0.36)	(0.14)	(0.21)	(0.21)	(0.83)
Spillovers:									0.03	-0.07
									(0.03)	(0.05)
Observations	13,648	6,824	13,648	6,824	13,648	6,824	13,648	6,824	13,648	6,824
Firm covariates	Yes	Yes	Yes	Yes	Yes	Yes			Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes
Panel B Implied Parameters										
Long-run Effects	-1.89	-1.23	-2.43	-1.68	-0.12	-0.94	-1.87	-2.34	-1.86	-1.32
$\hat{\delta}$	0.19	0.14	0.18	0.14	0.42	0.49	0.19	0.14	0.19	0.14
$\hat{\sigma}$	2.73	2.53	3.07	2.92	1.12	1.49	2.72	3.41	2.71	2.61
$\hat{\varepsilon}$	7.27	6.56	8.51	7.99	1.44	2.78	7.24	9.75	7.21	6.86
$\hat{\delta}\hat{\sigma}$	0.53	0.36	0.56	0.41	0.47	0.73	0.52	0.47	0.52	0.37
Panel C Equilibrium and Welfare Comparison										
n^{SP}/n^{EQ}	0.44	0.57	0.37	0.50	0.74	0.46	0.45	0.40	0.45	0.56
$U^{SP} - U^{EQ}$	0.03	0.02	0.04	0.03	0.01	0.03	0.03	0.04	0.03	0.02
Panel D Equilibrium and Welfare Comparison (Using Global SCC)										
n^{SP}/n^{EQ}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$U^{SP} - U^{EQ}$	13.74	8.94	16.99	11.60	3.50	12.15	13.62	15.58	13.55	9.50

Notes: This table presents regression estimates (Panel A), implied parameter values (Panel B) and implied distortions and welfare results (Panels C and D) for the health application. Regressions are unweighted and estimated with ordinary least squares and heteroscedasticity-robust standard errors clustered at firm level are presented in parentheses. Observations are at the firm-period level. The dependent variable is relative number of newly granted patents for dirty technologies relative to newly granted patents for clean technologies. Shifters are firm-level fuel prices adjusted (based on firm-level fuel consumption) inclusive of taxes (in $\ln(1+x)$ form unless otherwise indicated). All specifications include firm and period fixed effects as well as the values of government R&D subsidies for clean innovation, regulations over emissions, the relevant country's GDP per capita for that period (as in Aghion et al., 2016). Columns 1 and 2 replicate the main specifications from Table 2. Columns 3 and 4 use the inverse hyperbolic sine transformation. Columns 5 and 6 drop the firm fixed effects. Columns 7 and 8 additionally include the relative stock of knowledge in other firms in the same country as in the baseline specification of Aghion et al. (2016). In this case, we use the equations with spillovers for computing δ and σ. Panel C uses an externality number based on from Rennert et al.'s (2022) estimate of the social cost of CO2, converted to US-equivalent damages (see text for details), and Panel D uses their estimate for worldwide damages.

Appendix C: Data Sources and Construction

In this part of the Appendix, I describe the data sources for the three empirical exercises and provide some additional details.

Automation

Data for automation patents by country are directly from Acemoglu and Restrepo (2022). The flow of automation and non-automation patents were computed from the patents by different countries filed at the USPTO. In particular, all patents that are in the USPTO 901 class (technologies related to industrial robots) and all patents referencing this class are classified as automation patents. Aging variables are from the United Nations data, while the country-level covariates (GDP per capita, population, and average years of schooling) are from version 9.0 of the Penn World Tables (Feenstra, Inklaar and Timmer, 2015). Regressions are weighted by manufacturing value added in 1990 (sourced from the United Nations Industrial Development Organization).

Using these definitions, the exact estimating equations for columns 1 and 2 of Table 2 are:

$$\ln\left(\frac{\Delta n_{ct}}{n_{ct}}\right) = -\rho \ln n_{ct} + \chi \ln z_{ct} + \lambda \Delta \ln z_{ct} + \mathbf{X}_{c,1990}\boldsymbol{\gamma}_t + \epsilon_{ct}, \tag{C1}$$

where Δn_{ct} is the ratio of the flow relative automation patents (compared to non-automation patents) and n_{ct} is the relative technology stock (automation patent stock relative to non-automation patents stock). Stocks are computed from the corresponding flow variables using a 20% depreciation rate, as explained in the text.

The forcing variables are: the (log of) the ratio workers aged 56 and above to those between the ages of 25 and 55 and the (log of) 15- or 20-year ahead change in the ratio workers age 56 and above to those between the ages of 25 and 55. Finally, $\mathbf{X}_{c,1990}$ denotes the country covariates (log GDP per capita, log population, and average years of schooling, all in 1990), and the fact that its coefficient is time varying designates that these covariates are allowed to have a separate effect in every time period. The sample covers 69 countries and the time period 1986-2015. The observations are weighted by value added in manufacturing in 1990 and standard errors are clustered by country. The models estimated in Table B1 are variations of these equations as explained in the text.

In addition, the estimates in columns 11 and 12 of Table B1 are provided directly by Dechezleprêtre et al. (2022), based on their firm-level data set on automation and non-automation patents. The reader is referred to their paper for variable definitions and sources.

Medical Research

Our estimates for medical research's responsiveness to disease burden, depicted in columns 3 and 4 of Table 2 and Table B2, come directly from Acemoglu, Moscona, Sastry and Williams (2023). The estimating equation is similar to (24) in the text:

$$\ln\left(\frac{\Delta N_{dct}}{N_{dct}}\right) = \eta_d + \Gamma_c + \Upsilon_t - \frac{\rho}{2} \ln N_{dct} + \chi \ln Z_{dct} + \lambda \Delta \ln Z_{dct} + \epsilon_{dct}, \tag{C2}$$

where N_{dct} (ΔN_{dct}) is the stock (flow) of medical scientific articles on disease d in country c at time t. Stocks are again computed from flows using a 20% depreciation rate. The forcing variables are the level and change of disease burdens, defined as declines in the number of disability-adjusted life years caused

by a disease in a country and time period in our sample. These calculations are based on data from the Global Burden of Disease (GBD) project. Finally, η_d, Γ_c and Υ_t are, respectively, disease, country and time fixed effects, and in some specifications, two-way fixed effects are also included. All regressions in this case are unweighted. Additional details can be obtained from Acemoglu et al. (2023).

Energy

Data on the relationship between fuel prices and automobile patents come directly from Aghion et al. (2016). The data on flows of patents are based on the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office (EPO), and innovation is measured using a count of patents by application/filing date. The authors use data on tax-inclusive fuel prices, from the International Energy Agency (IEA), to compute a time-varying, country-specific fuel price by averaging the prices of diesel and gasolinespace prices. Country-specific fuel prices are then used to construct firm-level fuel prices as a weighted average of fuel prices across countries based on the firm's expected market share across countries (in practice, using a time-invariant share of the firm's sales in each market). The covariates are log GDP per capital (sourced from World Development Indicators), log R&D subsidies (from the IEA), and exposure to air pollution regulations. Emission regulations are for maximum level of tailpipe emissions for pollutants for new automobiles, coded between 0 and 5, and are taken from Dechezleprêtre et al. (2012).

The exact estimating equation for columns 5 and 6 of Table 2 is similar to but a little different from (C1). In particular, Aghion et al. (2016) impute log patent stocks as zero when stocks are zero. We avoid this by using $\ln(1+x)$ consistently for both flow and stock variables throughout this application. This gives our estimating equation as:

$$\ln\left(\frac{\Delta\tilde{n}_{fct}}{\tilde{n}_{fct}}\right) = -\rho\ln\left(\tilde{n}_{fct}\right) + \chi\ln z_{fct} + \lambda\Delta\ln z_{fct} + \mathbf{X}_{fct}\boldsymbol{\gamma} + \epsilon_{ct}, \tag{C3}$$

where

$$\ln\left(\frac{\Delta\tilde{n}_{fct}}{\tilde{n}_{fct}}\right) = \ln\left(\frac{1+\text{Patent}_{fct}^{\text{clean}}}{1+\text{Patent}_{fct}^{\text{dirty}}}\right) - \ln\left(\frac{1+\text{Stock}_{fct}^{\text{clean}}}{1+\text{Stock}_{fct}^{\text{dirty}}}\right)$$

and likewise,

$$\ln\left(\tilde{n}_{fct}\right) = \ln\left(\frac{1+\text{Stock}_{fct}^{\text{clean}}}{1+\text{Stock}_{fct}^{\text{dirty}}}\right),$$

with $\text{Patent}_{fct}^{\text{clean}}$ and $\text{Patent}_{fct}^{\text{dirty}}$, respectively, denoting the flow of clean and dirty automobile patents for firm f located in country c at time t, and $\text{Stock}_{fct}^{\text{clean}}$ and $\text{Stock}_{fct}^{\text{dirty}}$ likewise denoting the stocks of clean and dirty patents. The forcing variables, as described above, are based on firm-level fuel prices and their changes, while covariates are now time-varying but have constant coefficients. Regressions are unweighted and estimated by ordinary least squares, and standard errors are heteroscedasticity-robust and clustered at the country level. The models estimated in Table B3 are variations of these equations as explained in the text.

Markup Estimation

In this part of the Appendix, I describe our markup estimation strategies. Throughout, each firm is assumed to have a single, well-defined price at each point in time. Then, the gross markup of firm i at

time t is defined as

$$\Lambda_{it} = \frac{P_{it}}{MC_{it}}, \tag{C4}$$

where P_{it} is this firm's price at time t and MC_{it} is its marginal cost. Note that in the text I focused on net markups defined as

$$\mu_{it} = \Lambda_{it} - 1.$$

Production Function Estimation Methods

The production function method follows De Loecker et al. (2020). Let us first focus on a single industry, and suppose that each firm i in this industry has a production function

$$Q_{it}(V_{it}, K_{it}) \tag{C5}$$

at time t, with V_{it} denoting a composite of variable inputs (labor and material) and K_{it} representing its capital stock. Suppose that the capital stock is a quasi-fixed factor, meaning that it is chosen in advance (and hence the designation of the other factors as "variable"). The function Q_{it} is firm and time-varying, for example, it includes information on the firm's (revenue) productivity upon which variable costs may depend. In the estimation, the function Q_{it} will be taken to be Cobb-Douglas.

Consider the elasticity of this production function with respect to variable inputs, V_{it}, denoted by θ^V_{it}:

$$\theta^V_{it} = \frac{\partial Q_{it}}{\partial V_{it}} \frac{V_{it}}{Q_{it}} = \frac{1}{MC_{it}} \frac{P^V_{it} V_{it}}{Q_{it}},$$

where P^V_{it} is the price of the composite variable input, and the second equality exploits the fact that, because K_{it} is fixed, the marginal cost of production is $MC_{it} = P^V_{it} / (\partial Q_{it} / \partial V_{it})$. Next, using the definition of the markup in (C4) to substitute out MC_{it} and rearranging, we obtain

$$\Lambda_{it} = \theta^V_{it} \frac{P_{it} Q_{it}}{P^V_{it} V_{it}}.$$

Given this equation, firm-level markups can be estimated with data on revenue, $P_{it}Q_{it}$, cost of variable inputs, $P^V_{it} V_{it}$, and crucially the elasticity of the firm's production function with respect to variable inputs, θ^V_{it}.

Here I briefly outline De Loecker et al.'s (2020) estimation strategy, which I follow. Recall that the capital stock is quasi-fixed. Suppose also that observed sales are given by $\text{Sales}_{it} = \varepsilon_{it} Q_{it}(V_{it}, K_{it})$, where ε_{it} is a demand shifter realized after all input decisions are made. Finally, as noted below, suppose that the function Q_{it} is Cobb-Douglas, and denote the Hicks-neutral productivity of firm i at time t by Ω_{it}. Then we have

$$\ln \text{Sales}_{it} = \theta^V_t \ln V_{it} + \theta^K_t \ln K_{it} + \ln \Omega_{it} + \varepsilon_{it}, \tag{C6}$$

which allows for the Cobb-Douglas exponents, and thus output elasticities, to be time-varying, but constant across firms (within the industry being considered). The difficulty in the estimation of (C6) is that the firm knows Ω_{it} when choosing its composite variable input V_{it}, and thus OLS estimation will lead to biased output elasticities. De Loecker et al. (2020) deal with this problem by using a control function approach based on Olley and Pakes (1996). For example, Hicks-neutral productivity Ω_{it} can be assumed to be measurable with respect to the firm's capital stock K_{it}, investment I_{it}, and additional control variables

related to factor demands denoted by Z_{it}. This implies a relationship of the form

$$\ln(\Omega_{it}) = \phi_t(\ln K_{it}, \ln I_{it}, Z_{it}),$$

so that the elasticity of output with respect to variable inputs, θ_t^V, can be estimated from the following equation:

$$\ln \text{Sales}_{it} = \theta_t^V \ln V_{it} + \theta_t^K \ln K_{it} + \phi_t(\ln K_{it}, \ln I_{it}, Z_{it}) + \varepsilon_{it}. \tag{C7}$$

I follow De Loecker et al. (2020) and include the following terms in the ϕ function: a quadratic and cubic in $\ln K_{it}$, a main, quadratic and cubic in $\ln I_{it}$, and the interaction between these two variables, $\ln K_{it} \ln I_{it}$.[28] In addition, as in their specification, the Z_{it} variable includes the ratio of the firm's total costs to the four-digit industry total cost, and the ratio of the firm's total costs to the economy-wide total cost.

Once estimates of the variable input elasticity $\hat{\theta}_t^V$ are obtained, (gross) markups can be computed as

$$\hat{\Lambda}_{it}^P = \hat{\theta}_t^V \frac{P_{it} Q_{it}}{P_{it}^V V_{it}}, \tag{C8}$$

where the superscript P specifies that this is a markup estimated using the production function method. The (net) markup is then $\hat{\mu}_{it}^P = \hat{\Lambda}_{it}^P - 1$.

De Loecker et al.'s (2020) baseline estimates are based on a variant based on Ackerberg, Caves and Frazer (2015), where the composite variable input is used instead of investment and $\ln(\Omega_{it})$ is assumed to follow a first-order Markov process. For this specification, I directly use their estimates of these elasticities, reported in De Loecker et al. (2020a).

One drawback of the production function method is that the estimation of θ_{it}^V requires the model and the measurability assumptions embedded in the control function to be correctly specified.

As an alternative, De Loecker et al. (2020) also use cost shares to estimate θ_t^V. In particular, they compute industry-level output elasticities as

$$\tilde{\theta}_t^V = \text{median} \left\{ \frac{P_{it}^V V_{it}}{P_{it}^V V_{it} + R_t K_{it}} \right\},$$

where the median is across all firms within a two-digit industry, and R_t is the user cost of capital. In this case, (gross) markups can be obtained as

$$\hat{\Lambda}_{it}^C = \tilde{\theta}_t^V \frac{P_{it} Q_{it}}{P_{it}^V V_{it}},$$

where the superscript C refers to the fact that output elasticities are now are estimated from cost shares (but still taken to be common across firms within an industry). Then, naturally, $\hat{\mu}_{it}^C = \hat{\Lambda}_{it}^C - 1$. In practice, this approach gives similar results to the production function estimation.

One drawback in this case is that, although the functional form assumptions of the production function method are relaxed, the assumption that there is a common θ_t^V at the industry level is challenged by the fact that there is a large variation in cost shares, and the median is an arbitrary way of resolving this issue.

Another drawback of both approaches from my point of view is that it is not entirely clear whether markups relative to variable costs is the right notion for μ in the model, since this parameter captures

[28] As in their paper, investment in Compustat is computed from the capital stock data assuming a 10% depreciation rate, that is, $I_{it} = K_{it} - 0.9 \cdot K_{it-1}$.

how profitable a technology is and regulates incentives for innovation and entry. If one technology is more capital intensive and has higher markup estimates using the production function methods used by De Loecker et al. (2020), it may nonetheless have lower profitability and lower μ in the sense of the theoretical framework in the text. This motivates the next strategy.

The Accounting Method

As an alternative, one could directly use an estimate of profits to compute markups.

Using the same notation and terminology, (gross) markups are estimated in this case as

$$\hat{\Lambda}_{it}^{A} = \frac{P_{it}Q_{it}}{P_{it}^{V}V_{it} + R_{t}K_{it}}, \tag{C9}$$

where R_t is the user cost of capital (again assumed to be the same across firms in the industry). We then have: $\hat{\mu}_{it}^{A} = \hat{\Lambda}_{it}^{A} - 1$.

The drawback of this approach is well known: accounting profits do not correspond to economic profits. The advantage, on the other hand, is related to the discussion at the end of the previous subsection. This method takes into account capital costs explicitly, and thus may be more informative about the overall profitability of a technology/subsector.

Data

I follow De Loecker et al. (2020) and use Compustat North America for firm-level markup estimation. Compustat Fundamentals Annual extract is obtained through Wharton Research Data Service (WRDS), and I use the same variables as De Loecker et al. (2020a). Namely, the variable *SALE* measures revenues and variable costs are measured using the variable *COGS* (cost of goods sold, which includes expenses for materials, labor, overhead and other intermediate inputs). The capital stock of each firm is measured using the variable *PPEGT* (property, plant, and equipment gross total). The user cost of capital is also computed as in their paper: R_t = nominal interest rate$_t$−inflation$_t$+depreciation rate.[29, 30, 31] I set the depreciation rate at 10%. We exclude firms in the top and bottom 1% of cost of goods to sales ratio (*COGS/SALE*) and cost-shares, which are likely to have extreme values due to measurement error.[32]

Aggregating Markups

Throughout, I aggregate firm-level markups to industry-group level (in this instance, preventative health care vs. curative health care) by using the ratio of firm costs to industry-group costs. Specifically, our main estimates aggregate (gross) markups with the following equation:

$$\Lambda_{Jt}^{P} = \sum_{i \in J} \frac{\text{Cost}_{it}}{\text{Cost}_{jt}} \times \Lambda_{it}^{P},$$

[29] I follow De Loecker et al. (2020) and use the federal funds rate, FEDFUNDS, and the annual percent change in the relative price of investment goods, PIRIC. Both variables are taken from from the Federal Reserve Economic Data, FRED.

[30] Board of Governors of the Federal Reserve System (US), Federal Funds Effective Rate [FEDFUNDS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/FEDFUNDS, March 26, 2023.

[31] Relative Price of Investment Goods [PIRIC], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PIRIC, March 27, 2023.

[32] In particular, for this exercise, cost shares are measured as $\frac{COGS}{COGS+KEXP}$ and $\frac{COGS}{COGS+KEXP+SGA}$, where *SGA* measures selling, general, and administrative expenses

where $\text{Cost}_{it} = P_{it}^V V_{it} + R_t K_{it}$, and J denotes the industry-group in question (with $i \in J$ designating that firm i belongs to this group) such that $\text{Cost}_{jt} = P_{jt}^V V_{jt} + R_t K_{jt}$. I choose cost-based aggregation rather than using revenue-weights as in De Loecker et al. (2020), since, as these authors also recognize, revenue-based estimation can lead to inflated aggregate markups because high markup firms, which generate higher revenues, receive greater weights.

Firm Classification

This subsection explains how firms in Compustat are assigned to preventative and curative health care.

The classification is on the basis of the main North American Standard Industry Code (NAICS) assigned to firms in Compustat.[33]

Preventative: Health care firms whose main activity is in basic health provision, diagnosis or manufacture of vaccines and related products are assigned to the preventative health care group.[34] Firms with the following main NAICS codes are included in this category:

- *NAICS 621 - Ambulatory health services:* Firms that provide health care services diresctly or indirectly to to ambulatory patients and do not usually provide inpatient services. Includes outpatient services provided by physicians, dentists, and other health practitioners. Also includes outpatient care centers, medical and diagnostic laboratories, home health care services, and other ambulatory health care services.

- *NAICS 325413 - In-vitro diagnostic substances manufacturing:* Firms that manufacture in-vitro (i.e., not taken internally) diagnostic substances (chemical, biological, or radioactive substances). Substances are used for diagnostic tests, such as blood glucose , HIV, pregnancy, and other tests. It also involves manufacturing hematology, hormone, microbiology, and viral diagnostic substances, among others.

- *NAICS 325414 - Biological product (except diagnostic) manufacturing:* Firms primarily involved in manufacturing vaccines, toxoids, blood fractions, etc.

Curative: Health care firms whose main activity is in pharmaceutical preparation and high-tech medical equipment manufacturing firms (including a few that are related to advanced diagnostics) are assigned to the curative health care group. These firms are again identified based on their main NAICS codes, including the following categories:

- *325412 - Pharmaceutical preparation manufacturing:* Firms manufacturing in-vivo diagnostic substances and pharmaceutical preparations (except biological) intended for internal and external consumption in dose forms, such as tablets, capsules, vials, ointments, powders, solutions, and suspensions.

- *334510 - Electromedical and electrotherapeutic apparatus manufacturing:* Firms manufacturing electromedical and electrotherapeutic apparatus such as magnetic resonance imaging equipment, medical ultrasound equipment, pacemakers, hearing aids, electrocardiographs, and electromedical endoscopic equipment.

[33] Codes and descriptions obtained from the U.S. Census Bureau North American Industry Classification System (NAICS) at https://www.census.gov/naics/?99967.

[34] In addition, preventative health care should also include those in the area of social assistance, *NAICS 624*, which comprises firms providing individual and family services, community food and housing, vocational rehabilitation services, child daycare services, as well as emergency and other relief services. Nevertheless, there are no firms in this NAICS category in Compustat.

- *339112 - Surgical and medical instrument manufacturing:* Firms manufacturing medical, surgical, ophthalmic, and veterinary instruments and apparatus (except electrotherapeutic, electromedical, and irradiation apparatus). Examples are syringes, needles, anesthesia apparatus, blood transfusion equipment, catheters, surgical clamps, and medical thermometers.

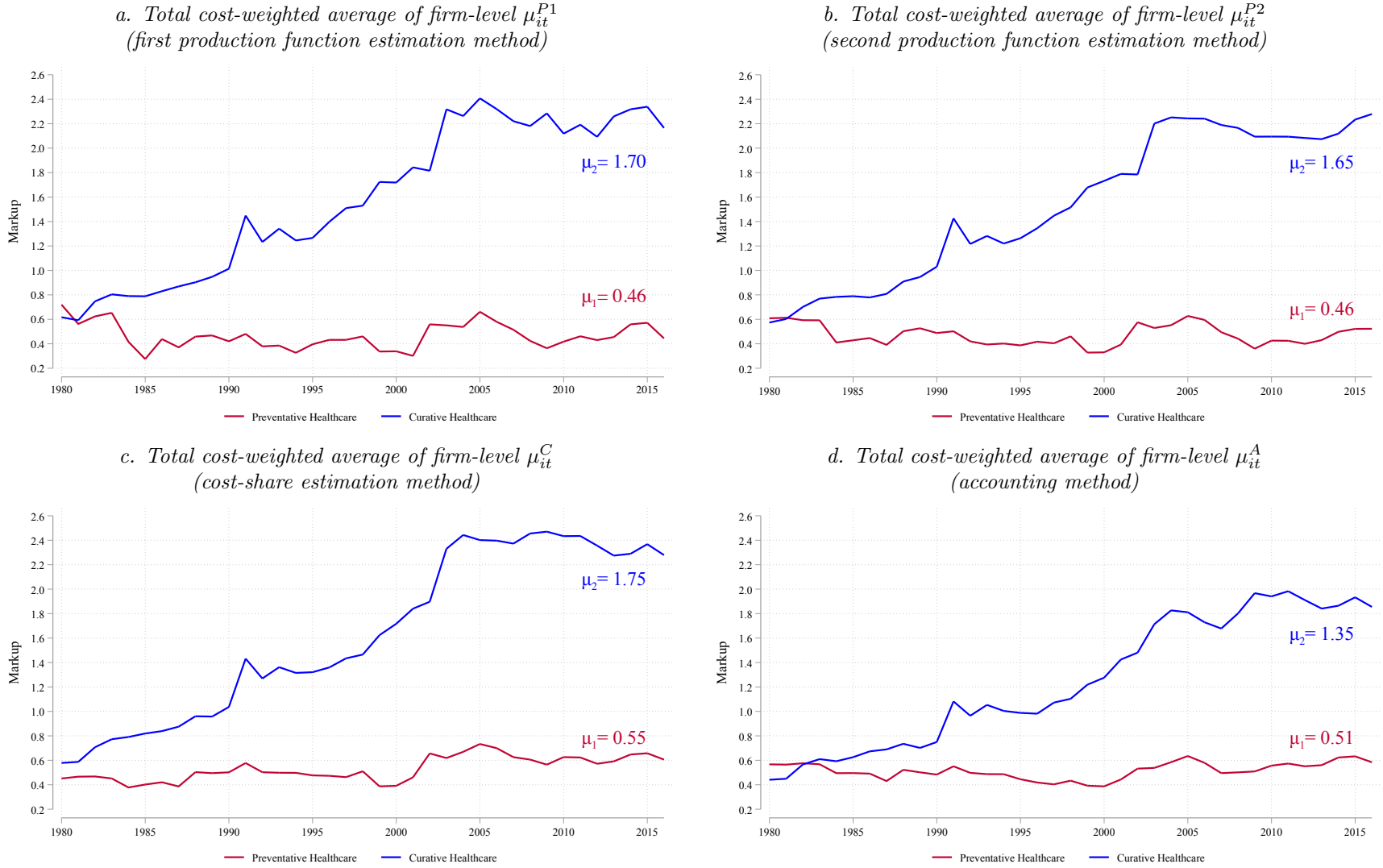
- *339113 - Surgical applicance and supplies manufacturing:* Firms manufacturing surgical appliances and supplies such as orthopedic devices, prosthetic appliances, surgical dressings, personal safety equipment, hospital beds, operating tables, etc.

Table D1, included in Appendix D, which is available upon request, provides a full list of health care firms in Compustat and their assignment to preventative and curative categories. It also lists the relevant sample period for the firm, sales, costs of goods sold, and capital stock, as well as the four measures of (net) markup—based on production function, cost share and accounting methods. In total, our sample includes 658 preventative and 1,069 curative health care firms. At the bottom of each panel, (cost-weighted) averages of the markups are also presented.

The four panels of Figure C1 show the evolution of markups based on the two production function estimation methods, cost share and accounting methods, separately for firms in the preventative and curative categories. Each panel also gives the average (net) markup, which corresponds to μ in the model. The trends are fairly similar with the different methods and show some fluctuations and also a significant increase in markups among firms in the curative category. This is consistent with the patterns reported by De Loecker et al. (2020) at a higher level of aggregation. The increase in markups among curative firms is in fact larger than at the two-digit level patterns depicted by De Loecker et al. (2020).

Our baseline uses the averages in Panel a, which give $\mu_1 = 1.70$ for preventative firms and $\mu_2 = 0.46$ for curative firms as also shown in Table 1 in the text. The numbers in the other panels are quite similar, and using these numbers instead yields broadly similar results to those reported in Table 2. Table C1 shows the technology ratio and welfare loss estimates corresponding to the specifications in Table B2, if instead we use the markup estimates in Panels b, c or d of Figure C1.

Figure C1: Aggregate Markups by Sector Group



Note: This figure depicts total cost-weighted averages of firm-level markups across the preventative and curative technology groups. Cost shares are defined as $(COGS_{it} + R_t PPEGT_{it})/(COGS_{jt} + R_t PPEGT_{jt})$ (see Appendix D Table D1). The four panels use firm-level markups $\mu_{it}^{P}1$, $\mu_{it}^{P}2$, μ_{it}^{C} and μ_{it}^{A}, which are based, respectively, on the first and second production function estimation methods, cost-share estimation method and the accounting method, as described in Section 6. The list of firms is given in Appendix D Table D1.

Table C1: Sensitivity Analysis of Technology Distortions and Welfare Losses from Markups

LHS	$\ln(x)$	$\ln(x)$	$\ln(1+x)$	$\ln(1+x)$	$\text{asinh}(x)$	$\text{asinh}(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$	$\ln(x)$
Samples	Full	Full	Full	Full	Full	Full	Full	Full	Full	Full	US	US
Frequency	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year	5-year	10-year
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A Parameters Estimated from Regressions												
Initial Relative Stock: $\hat{\rho}$	0.93	1.11	0.36	0.83	0.38	0.84	0.45	0.48	0.93	1.09	0.94	1.27
	(0.03)	(0.03)	(0.02)	(0.04)	(0.02)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.10)	(0.12)
Initial Shifter: $\hat{\chi}$	0.10	0.14	0.05	0.11	0.07	0.13	0.07	0.10	0.11	0.14	0.32	0.26
	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.12)	(0.10)
Changes in Shifter: $\hat{\lambda}$	-0.004	0.001	-0.01	0.002	-0.01	0.002	-0.04	-0.06	-0.004	0.01	0.26	0.12
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.17)	(0.07)
Observations	55,699	37,389	75,399	44,569	75,399	44,569	55,702	37,394	55,625	37,358	1,243	741
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes						
Disease fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes			Yes	Yes
Period-Country fixed effects									Yes	Yes		
Period-Disease fixed effects									Yes	Yes		
Panel B Implied Parameters												
Long-run Effects	0.11	0.14	0.15	0.13	0.18	0.16	0.16	0.22	0.11	0.14	0.34	0.26
$\hat{\delta}$	0.07	0.00	0.64	0.17	0.62	0.16	0.55	0.52	0.07	0.00	0.06	0.00
$\hat{\sigma}$	1.10	1.14	1.05	1.11	1.06	1.13	1.07	1.09	1.11	1.14	1.31	1.26
$\hat{\varepsilon}$	1.18	1.26	1.09	1.19	1.11	1.24	1.12	1.17	1.19	1.26	1.57	1.47
$\hat{\delta}\hat{\sigma}$	0.08	0.00	0.67	0.19	0.66	0.18	0.58	0.57	0.08	0.00	0.08	0.00
Panel C Equilibrium and Welfare Comparison (Markups Only), Baseline												
n^{SP}/n^{EQ}	0.43	0.45	0.11	0.39	0.11	0.38	0.17	0.17	0.43	0.45	0.37	0.42
$U^{SP} - U^{EQ}$	0.06	0.05	0.15	0.06	0.15	0.07	0.12	0.12	0.06	0.05	0.06	0.06
Panel D Equilibrium and Welfare Comparison (Markups Only), Production function estimation method 2												
n^{SP}/n^{EQ}	0.44	0.46	0.12	0.40	0.12	0.39	0.18	0.18	0.44	0.46	0.38	0.43
$U^{SP} - U^{EQ}$	0.05	0.05	0.14	0.06	0.14	0.06	0.11	0.11	0.05	0.05	0.06	0.05
Panel E Equilibrium and Welfare Comparison (Markups Only), Cost-share based estimation												
n^{SP}/n^{EQ}	0.50	0.51	0.16	0.45	0.16	0.45	0.23	0.23	0.50	0.51	0.44	0.48
$U^{SP} - U^{EQ}$	0.04	0.04	0.10	0.04	0.10	0.05	0.08	0.08	0.04	0.04	0.05	0.04
Panel F Equilibrium and Welfare Comparison (Markups Only), Accounting method												
n^{SP}/n^{EQ}	0.53	0.55	0.19	0.49	0.20	0.48	0.26	0.27	0.53	0.55	0.47	0.52
$U^{SP} - U^{EQ}$	0.03	0.03	0.08	0.04	0.08	0.04	0.07	0.07	0.03	0.03	0.04	0.03

Notes: This table shows how technology distortions and welfare losses change across the specifications considered in Table B2 for different values of markups. Panels A and B replicate the same panels in Table B2. The remaining four panels correspond to the four sets of markup estimates, $\hat{\mu}_{it}^{P1}$, $\hat{\mu}_{it}^{P2}$, $\hat{\mu}_{it}^{C}$ and $\hat{\mu}_{it}^{A}$, which are, respectively, from the first and second production function estimation methods, cost-share estimation method and the accounting method. See Figure C1 and Section 6 on the markup estimates, and see Table B2 on the parameter estimates and the underlying regression models for the different columns here.

Quality-Adjusted Life Years

In this section, we describe how differences between preventative and curative technologies in terms of quality-adjusted life years (QALYs) are estimated.

Methodology

Quality-Adjusted Life Years (QALYs) are a common measure used for evaluating the effectiveness of medical treatments and interventions. They quantify the overall gains in quantity and quality of life. QALYs are calculated by multiplying the number of years of life gained by a quality of life scale, which ranges from 0 (death) to 1 (perfect health). To access cost-effectiveness analyses in a comprehensive manner, we use the Cost-Effectiveness Analysis (CEA) Registry by the Center for the Evaluation of Value and Risk in Health, Tufts Medical Center. This registry includes studies on a wide range of health interventions, including drugs, medical devices, diagnostic tests, and prevention strategies and reports detailed information on the methods used in and results of each study.

We restrict the sample to modern healthcare innovations with studies conducted in the United States and benchmark the relevant innovation to the year of Food and Drug Administration (FDA) approval. We exclude a large number of studies included in the registry that evaluate the effectiveness of immunization drives and information campaigns. We focus on studies on pharmaceuticals, medical devices, and surgical procedures, especially those that compare a drug to placebo or no treatment. For these innovations we extract the QALYs gained per patient from the relevant journal article or website containing the study. In the case that a drug is compared to another drug instead of a placebo or no treatment, where possible, we search for auxiliary studies that compare one of the drugs in the main study to placebo and use that as a reference point to impute the effect and cost of all drugs in the main study relative to placebo.[35] Note that the QALY numbers obtained from this procedure can be negative, if new procedures are worse than no treatment or placebo, and are indeed so in a few cases.

Most estimates give QALY gains per patient. To construct comparable social benefits, I convert these estimates into QALY gains per dollar. Specifically, I use the following equation for each innovation i:

$$\text{QALY per dollar}_i = \frac{\text{QALY per patient}_i \times \text{Number of users}_i}{\text{Cost per user}_i \times \text{Number of users}_i + \text{R\&D costs}_i}. \tag{C10}$$

Intuitively, this expression corresponds to total benefits divided by total costs, including R&D costs. In estimating the number of users, I limit the horizon for each innovation to 20 years, which amounts to assuming that this innovation will be replaced by a new one on average every 20 years. Given these estimates, I construct the average quality-adjusted life year gains by preventative and curative technology groups as

$$\text{QALY per dollar}_G = \sum_{i \in G} \text{Cost share}_i \times \text{QALY per dollar}_i,$$

where G is either the preventative or the curative technology group, and

$$\text{Cost share}_i = \frac{\text{Cost per user}_i \times \text{Number of users}_i + \text{R\&D costs}_i}{\sum_{i' \in G} \left(\text{Cost per user}_{i'} \times \text{Number of users}_{i'} + \text{R\&D costs}_{i'}\right)}.$$

In these equations, R&D costs are estimated from the medical literature, which provides average of

[35] In principle, one might wish to obtain QALYs relative to a single dominant treatment that exists before the innovation. In practice, this did not prove to be straightforward, and hence I opted for making all comparisons relative to placebo or no treatment.

R&D costs by class of drugs, e.g., oncological, immunomodulant, therapeutic recombinant proteins and mAbs, cardiovascular, etc. The medical papers in this literature use a variety of methods to obtain R&D costs, including using proprietary databases with cost information at the individual drug level, mandatory SEC filings, and industry surveys. Virtually all papers involve accounting for both failed and approved drugs, the type and duration of clinical trials, and the status of drug review at the FDA such as fast track, accelerated approval, or priority review. These papers are helpfully reviewed in Table 1 of Schlander et al. (2021). We match each innovation to its pharmaceutical category and impute the cost of R&D as the average cost for that group of drugs. For example, amlodipine is a calcium channel blocker that can treat high blood pressure and chest pain. As it acts on the cardiovascular system, we impute its R&D cost as the average R&D cost for all cardiovascular drugs. For surgical procedures, we use the sum of R&D expenses over several years or total invested capital for the primary manufacturer of equipment used in the surgical procedure, whichever data are available.

Per-patient usage costs are taken from the same papers that present the QALY benefits. These costs are often constructed as a sum of a direct treatment cost and an indirect health care cost, which imputes a production loss due to the patient's injury and inability to work. Note that both the QALY estimates and per-patient usage cost are relative to placebo or no treatment, and thus we obtain negative values in a few cases. This is primarily because treatment avoids other costs patients incur in the future.

We use three methodologies to estimate the number of users. First, we look for direct estimates of the number of patients using the drug. Such statistics are available on clincalc.com or the Centers for Disease Control and Prevention (CDC) website. Second, if no direct estimate is available, we estimate the number of users by dividing the US total sales for the innovation by the annual therapy cost, which is itself the product of dosage, frequency, and price per dosage. Lastly, if the direct and revenue imputation approach are infeasible, we gauge the number of patients by multiplying the incidence of disease by the proportion of patients who undergo treatment by the innovation.

For one-time innovations, such as some surgical procedures, we focus on the number of annual patients and multiply this by 20. For innovations that involve recurring use, such as antihypertensive medication, the number of patients is given by the contemporaneous usage prevalence, under the assumption that a patient uses the drug for the duration of the time horizon.

Finally, given the QALY per dollar$_G$ estimates, we set $\tilde{\tau}_1 = 0$, and compute $\tilde{\tau}_2$ on the basis of the relative shortfall of the curative technologies compared to the preventative technologies:

$$\tilde{\tau}_2 = 1 - \frac{\text{QALY per dollar}_{curative}}{\text{QALY per dollar}_{preventative}}.$$

List of Procedures

I now provide further details on the procedures and innovations selected in the computation of the QALY numbers.

The CEA Registry contains roughly 10,000 entries for the United States with QALY outcomes. First, we manually considered each study to determine if it studied an innovation which constituted a modern healthcare innovation. This step eliminated immunization drives and information campaigns. Second, we excluded studies without a placebo or no treatment comparison. This step dropped a significant proportion of the sample, including studies comparing dosages of a given drug, evaluating different drugs for a given disease, or assessing the most efficacious treatment combination of a set of drugs (many of these studies are aimed at better informing *clinical* use, which is very different from our purpose here). Third, as screening procedures are not easily categorized into preventative vs. curative, we dropped all

such studies. Lastly, we sought the background and history of each innovation and kept those which were commercially developed in the late 20th century.

We then performed a second pass where we actively searched for cost-effectiveness studies relating to the top 20 drugs in the United States, as listed in clincalc.org. While data were not available for all 20 drugs, we were able to add 9 additional important drugs to our list.

Table D2 in Appendix D (available upon request) lists the 71 procedures we consider. In each case we provide reference to the source article where the medical information is taken, and list our estimates of R&D costs, usage costs, total QALY benefits, and our final QALY benefits per dollar. The two panels correspond to curative and preventative technologies.

Additional References for Appendix C

Ackerberg, Daniel, Kevin Caves and Garrett Frazer (2015) "Identification Properties of Recent Production Function Estimators." *Econometrica*, 83(6): 2411-2451.

Dechezleprêtre, Antoine, Richard Perkins, and Eric Neumayer (2012) "Regulatory Distance and the Transfer of New Environmentally Sound Technologies: Evidence from the Automobile Sector." Working Paper no. 2012.33, Fondazione Eni Enrico Mattei, Milan.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger (2020a) "Replication Data for: The Rise of Market Power and the Macroeconomic Implications." *The Quarterly Journal of Economics*, Harvard Dataverse, V1, https://doi.org/10.7910/DVN/5GH8XO.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (2015) "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150-82.

Olley G. Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunication Equipment Industry." *Econometrica*, 64(6): 1263-1297.

Schlander, M., Hernandez-Villafuerte, K., Cheng, C. Y., Mestre-Ferrandiz, J., & Baumann, M. (2021) "How Much Does It Cost to Research and Develop a New Drug? A Systematic Review and Assessment." *PharmacoEconomics*, 39, 1243-1269.